GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 22:10:58 ; Search time 876.5 Seconds
(without alignments)
6590.438 Million cell updates/sec

Title: US-10-664-859-1
Perfect score: 12037
Sequence: 1 acgagtgcttctcttattat.....aaaaaaaaaaaaaaaaaaaa 6909

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 3730428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10664859/runat_02112005_171806_3577/app_query.fasta_1.7047
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10664859_@CGN_1_1_982_@runat_02112005_171806_3577
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7040	58.5	1429	20	US-11-097-143-3129	Sequence 3129, Ap
2	570	4.7	112	9	US-09-915-543-10	Sequence 10, Appl
3	570	4.7	112	14	US-10-322-579-10	Sequence 10, Appl
4	570	4.7	112	15	US-10-664-859-10	Sequence 10, Appl
5	321.5	2.7	1435	15	US-10-276-774-2178	Sequence 2178, Ap
6	314	2.6	1394	15	US-10-381-247B-18	Sequence 18, Appl
7	314	2.6	1394	16	US-10-723-860-1764	Sequence 1764, Ap
8	314	2.6	1394	18	US-10-756-149-5163	Sequence 5163, Ap
9	312	2.6	1426	9	US-09-915-543-15	Sequence 15, Appl
10	312	2.6	1426	14	US-10-322-579-15	Sequence 15, Appl
11	312	2.6	1426	15	US-10-664-859-15	Sequence 15, Appl
12	312	2.6	1426	16	US-10-408-765A-1958	Sequence 1958, Ap
13	296.5	2.5	1450	15	US-10-381-247B-17	Sequence 17, Appl
14	296.5	2.5	1494	15	US-10-381-247B-2	Sequence 2, Appli
15	269.5	2.2	1594	20	US-11-097-143-17856	Sequence 17856, A
16	258.5	2.1	1366	20	US-11-097-143-27555	Sequence 27555, A
17	258.5	2.1	2151	20	US-11-097-143-7050	Sequence 7050, Ap
18	256.5	2.1	5322	17	US-10-732-923-8729	Sequence 8729, Ap
19	247.5	2.1	2285	18	US-10-773-446-101	Sequence 101, App
20	242.5	2.0	1161	20	US-11-097-143-38793	Sequence 38793, A
21	241.5	2.0	1778	20	US-11-097-143-7035	Sequence 7035, Ap
22	241	2.0	2280	20	US-11-097-143-11742	Sequence 11742, A
23	240	2.0	2112	20	US-11-097-143-8001	Sequence 8001, Ap
24	233.5	1.9	1243	16	US-10-408-765A-1964	Sequence 1964, Ap
25	233.5	1.9	1441	15	US-10-355-218-2	Sequence 2, Appli
26	233.5	1.9	1441	18	US-10-503-050A-2	Sequence 2, Appli
27	233	1.9	2703	20	US-11-097-143-16689	Sequence 16689, A
28	233	1.9	3190	17	US-10-732-923-18448	Sequence 18448, A
29	230.5	1.9	3275	18	US-10-840-060-90	Sequence 90, Appl
30	230.5	1.9	3275	20	US-11-097-143-38103	Sequence 38103, A
31	230	1.9	3276	17	US-10-732-923-18447	Sequence 18447, A
32	229.5	1.9	2005	9	US-09-735-367B-3	Sequence 3, Appli
33	228.5	1.9	2063	9	US-09-735-367B-2	Sequence 2, Appli
34	228.5	1.9	2063	18	US-10-973-858-8	Sequence 8, Appli
35	228.5	1.9	2063	18	US-10-450-763-47506	Sequence 47506, A
36	228	1.9	1424	15	US-10-418-027-1	Sequence 1, Appli
37	228	1.9	1424	16	US-10-473-127-347	Sequence 347, App
38	228	1.9	1424	16	US-10-473-127-350	Sequence 350, App
39	228	1.9	1424	16	US-10-473-127-360	Sequence 360, App
40	223.5	1.9	1522	16	US-10-473-127-355	Sequence 355, App
41	222.5	1.8	1441	14	US-10-170-682-3	Sequence 3, Appli
42	222.5	1.8	1459	15	US-10-369-493-6418	Sequence 6418, Ap
43	222.5	1.8	1966	16	US-10-408-765A-2217	Sequence 2217, Ap
44	222.5	1.8	2440	15	US-10-341-434-236	Sequence 236, App
45	222.5	1.8	2440	18	US-10-885-977-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-11-097-143-3129
; Sequence 3129, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769

```
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3129
;   LENGTH: 1429
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-3129

Alignment Scores:
Pred. No.:              0           Length:         1429
Score:                  7040.00     Matches:        1425
Percent Similarity:     71.09%      Conservative:   1
Best Local Similarity:  71.04%      Mismatches:     3
Query Match:            58.49%      Indels:         577
DB:                     20          Gaps:           6

US-10-664-859-1 (1-6909) x US-11-097-143-3129 (1-1429)

Qy        453 ATGCTCTCGACAACAATGCCCCGCAGTCCAACCCAACAACAGCCGCAACCAAACTCCGAT 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuSerThrThrMetProArgSerProThrGlnGlnGlnProGlnProAsnSerAsp 20

Qy        513 GCCTCCTCAACAAGTGCATCTGGATCAAATCCTGGAGCAGCGATCGGAAATGGGGACTCG 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaSerSerThrSerAlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSer 40

Qy        573 GCGGCGAGCAGAAGTTCTCCGAAGACCCTTAATAGCGAACCCTTTTCTACTTTGTCGCCG 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AlaAlaSerArgSerSerProLysThrLeuAsnSerGluProPheSerThrLeu-Ser-- 59

Qy        633 GGTAAGACTTGTATTGATTTCTCTTTGTCCGGAATTATAACAACTTTCTGTGTTTCCAGA 692
                                                                  |||||
Db         60 ------------------------------------------------------ProAs 61

Qy        693 TCAAATAAAATTGACGCCAGAAGAAGGCACTGAGAAAAGCGGACTATCAACTAGTGATAA 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 pGlnIleLysLeuThrProGluGluGlyThrGluLysSerGlyLeuSerThrSerAspLy 81

Qy        753 AGCTGCCACTGGAGGAGCCCCAGGCAGTGGAAATAATCTGCCCGAGGGACAAACTATGCT 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 sAlaAlaThrGlyGlyAlaProGlySerGlyAsnAsnLeuProGluGlyGlnThrMetLe 101

Qy        813 AAGGCAGAACTCTACGAGCACAATCAACTCGTGCCTAGTCGCTTCTCCACAAAACTCCAG 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 uArgGlnAsnSerThrSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSe 121

Qy        873 TGAACACTCGAATAGCAGCAATGTGTCTGCTACAGTGGGCCTTACTCAGATGGTAGATTG 932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 rGluHisSerAsnSerSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCy 141

Qy        933 TGACGAGCAATCGAAGAAAAACAAATGTAGTGTGAAGGACGAGGAAGCTGGTAAGACTGC 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 sAspGluGlnSerLysLysAsnLysCysSerValLysAspGluGluAla----------- 157

Qy        993 CCTACAAATGGTTTAAAATTTTAAAATGTATTGGCGTTCACCTTTGTTAATCATTTAATT 1052

Db        157 ------------------------------------------------------------ 157

Qy       1053 GTTTTTTTTTTGCTATACTTACAATTTTAGTTTTAAACTTGTAAACTTGACTAAAACTCG 1112

Db        157 ------------------------------------------------------------ 157

Qy       1113 CGAAGCTCGGATCAAAACAGACATTTTCTTGGAACCGTAATTAAGCTCATAAAAATATTA 1172

Db        157 ------------------------------------------------------------ 157

Qy       1173 ATTCATCTTGATGGAATGCATATCATAGATGTACTCAAACATCTCAAGAAAGACCTCAAA 1232

Db        157 ------------------------------------------------------------ 157

Qy       1233 TTGGATCAACTAATTAGTTTGAGAAAAAATTGCTGTACTTTTAAGAATATATTAATTTAA 1292

Db        157 ------------------------------------------------------------ 157

Qy       1293 AAATTTGCTGAGTGAAATGATATAATAGTCACAATAATTTTTAGTTAAACTGCTAAAGCA 1352

Db        157 ------------------------------------------------------------ 157

Qy       1353 TTTTGAATAGCCGTGCTACGCAGATGCTACTAGACGCGGTGTAAAAGCTAATTTTTATTT 1412

Db        157 ------------------------------------------------------------ 157

Qy       1413 AAAAGCTGTCCTAATATTCCATAACCATTAATGTCCCATTTCAGAAATAAGTTCTAATAA 1472
                                                     ||||||||||||||||
Db        158 ----------------------------------------GluIleSerSerAsnLy 163

Qy       1473 AGCAAAAGGTCAAGCAGCTGGTGGCGGCTGCGAAACAGGTTCTACATCCAGTTTGACTGT 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 sAlaLysGlyGlnAlaAlaGlyGlyGlyCysGluThrGlySerThrSerSerLeuThrVa 183

Qy       1533 CAAGGAAGAACCCACCGATGTCTTAGGCAGTTTAGTAAATATGAAAAAAGAAGAAAGAGA 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 lLysGluGluProThrAspValLeuGlySerLeuValAsnMetLysLysGluGluArgGl 203

Qy       1593 AAATCATTCGCCAACGATGTCCCCTGTTGGTTTTGGTTCAATTGGTAATGCACAGGACAA 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 uAsnHisSerProThrMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAs 223

Qy       1653 CTCCGCTACACCGGGTAAGTTTTAAGAGATCCATATAAAGCAAATAACAAGAATTAATGT 1712
              ||||||||||
Db        223 nSerAlaThr------------------------------------------------- 226

Qy       1713 CAGTTACCAATTTTATTTGATAGTCAAAGAACTACTATAGCGATATCTCCTGCCTTTTAA 1772

Db        226 ------------------------------------------------------------ 226

Qy       1773 TTTTATTTTAATTAGGAAATACGAATATTTCTAATTTGTAAAATAAAATTGATTAATTAA 1832

Db        226 ------------------------------------------------------------ 226

Qy       1833 CTAGAATTTAAAAACCTTTTGAATTAGGACATACCCTTCCAAAAATCAGTAATCATTGGG 1892

Db        226 ------------------------------------------------------------ 226

Qy       1893 AACGAGAGTGTGGTCCCGAAGGAGACTACTATAAAACCTTTTGAGCTATCTGATACTGCA 1952

Db        226 ------------------------------------------------------------ 226

Qy       1953 CGCTACTAAAAATGATTAGTTTAGGAAAATGGGTGTAATTTTGTAGGAAGTTTTCATTTT 2012

Db        226 ------------------------------------------------------------ 226

Qy       2013 AGAAGAAATGTGATTATTTTATTAAACCCCTTCAAGCGGAACTACATTTGTTCTACGATA 2072

Db        226 ------------------------------------------------------------ 226

Qy       2073 TTTTGGAAAAACAAATGGTTAAGTTGGAAAGTGCCTATAAAACAGAATTCCACGGTTTCA 2132

Db        226 ------------------------------------------------------------ 226

Qy       2133 AATACTAACCAGGTTTTTGATTTAATTTTGATTAAATGAGAAATTATCACACTTCAGTTA 2192

Db        226 ------------------------------------------------------------ 226

Qy       2193 AAATGTTTAATTCGATTAAGGTCGGACAATCACAGCAGATTTCCATTTTTGCGTGTATAT 2252

Db        226 ------------------------------------------------------------ 226
```

```
Qy       2253 ATAGAAGTCGCCTTCACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCCCGCA 2312
Db        226 ------------------------------------------------------------ 226
Qy       2313 GTGATTTATATAGATGATTTACGAGTTATTTAATTTTTTATGGTGTATTTTAATAAATAT 2372
Db        226 ------------------------------------------------------------ 226
Qy       2373 CTTATTTATTCATTTTACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAA 2432
                                                  ||||||||||||||||||||||||
Db        227 ----------------Pro--ValLysIleGluArgIleSerAsnAspSerThrThrGlu 240
Qy       2433 AAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAG 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LysLysGlySerSerLeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGln 260
Qy       2493 TTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTA 2552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 LeuAsnProAspPheIleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuVal 280
Qy       2553 AGCGGAGTAGGACCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACT 2612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 SerGlyValGlyProIleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThr 300
Qy       2613 GCCAACGCCAATGGAATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAA 2672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AlaAsnAlaAsnGlyIleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGln 320
Qy       2673 AATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGC 2732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AsnHisIlePheValPheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSer 340
Qy       2733 GGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTG 2792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlyGlnPheGlnThrIleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeu 360
Qy       2793 GAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTC 2852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GluAspPhePheMetLysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerVal 380
Qy       2853 GGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAA 2912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyMetProTrpIleGlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLys 400
Qy       2913 ATAACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAA 2972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 IleThrGlnGlnGlnProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGln 420
Qy       2973 CATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCT 3032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 HisGluAsnSerLysArgSerThrValSerAlaProSerAsnSerPheValAspGlnSer 440
Qy       3033 GATCCTATGGGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT 3092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AspProMetGlyAsnGluThrGluLeuMetCysTrpGluGlyGlySerSerAsnThrSer 460
Qy       3093 AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAG 3152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 ArgSerGlyGlnAsnSerArgAsnHisValAspSerIleSerThrSerSerGluSerGln 480
Qy       3153 GCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGAT 3212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLysIleLeuGluAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAsp 500
Qy       3213 CCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAA 3272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ProGlyLeuThrThrGluAsnAsnIleValSerLeuGlnGlyValLysValProAspGlu 520
Qy       3273 AACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATG 3332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 AsnLeuThrProGlnGlnArgGlnHisArgGluGluGlnLeuAlaLysIleLysLysMet 540

Qy       3333 AATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATA 3392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AsnGlnPheLeuPheProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIle 560
Qy       3393 ACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATA 3452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 ThrLysIleProGlyAspLeuMetMetGlyMetSerGlyGlyGlyGlyGlySerIleIle 580
Qy       3453 AATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCG 3512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 AsnProThrMetArgGlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAla 600
Qy       3513 ACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGT 3572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ThrSerSerGlyLeuSerGluAspValMetHisProGlyAspValIleSerAspMetGly 620
Qy       3573 GCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGT 3632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 AlaValIleGlyCysAsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGly 640
Qy       3633 GTTGTCACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGC 3692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGly 660
Qy       3693 GCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACA 3752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThr 680
Qy       3753 AGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGC 3812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 SerCysAsnValIleGlyThrAlaProAspMetSerLysGluValLeuAsnGlnAspSer 700
Qy       3813 CGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ArgThrHisSerHisGlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHisGln 720
Qy       3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 PhePheGluGluArgLeuLysGlyGlyLysProArgGlnValThrGlyThrValValPro 740
Qy       3933 CAACAGCAAACCCCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTG 3992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 GlnGlnGlnThrProSerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeu 760
Qy       3993 CAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACT 4052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 GlnGlyProProProProTyrHisSerIleGlnArgSerAlaSerValProIleAlaThr 780
Qy       4053 CAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCA 4112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThrAla 800
Qy       4113 GCAGTCATGGGATTGCCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGA 4172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 AlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGly 820
Qy       4173 TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCA 4232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 SerValProGlnAlaAsnThrSerThrValGlnAlaGlyThrThrThrValLeuSerAla 840
Qy       4233 AACAAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAAT 4292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AsnLysAsnCysPheGlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsn 860
Qy       4293 ACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCT 4352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 ThrGlySerSerSerValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSer 880
Qy       4353 CATCTATCCCCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTG 4412
              ||||||||||||||||||||||||||||||||||||
Db        881 HisLeuSerProLysGluPheGluSerPheGlyGlnSerSer---------------- 894
Qy       4413 TTTAATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACA 4472
```

where are nts 5482-5601 enc aa 1140-1179?

```
Db        894 ------------------------------------------------------------ 894
Qy       4473 TAACTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG 4532
Db        894 ------------------------------------------------------------ 894
Qy       4533 CCCATCCATTTCTTGAATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAA 4592
Db        894 ------------------------------------------------------------ 894
Qy       4593 TTAAACACGAAAATTCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAA 4652
Db        894 ------------------------------------------------------------ 894
Qy       4653 AATATGTTTATTATAATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCAC 4712
                                  ||||||||||||||||||||||||||||||||
Db        895 --------------------Ala--GlyAspAsnMetLysSerArgArgProSerProG 907
Qy       4713 AGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTG 4772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 lnGlyGlnArgSerProValAsnSerLeuIleGluAlaAsnLysAspValArgPheAlaA 927
Qy       4773 CATCCAGTCCTGGTTTTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAA 4832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 laSerSerProGlyPheAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuA 947
Qy       4833 ACGCCTATAAAATGGGCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATT 4892
              |||||||||||||||||||||||||||||||||||  :::
Db        947 snAlaTyrLysMetGlySerThrAsnIleGlnMetGlu--Arg----------------- 960
Qy       4893 TAACGTTTTTGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATC 4952
                                                  ||||||||||||||||||||||
Db        961 ------------------------------------GlnAlaSerAlaGlnGlyGlySe 968
Qy       4953 CGTACAATTTAGTCGGCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCC 5012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 rValGlnPheSerArgArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgPr 988
Qy       5013 GCCACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCA 5072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 oProProAsnLysMetThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGl 1008
Qy       5073 ACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGAT 5132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 nGlnLeuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMe 1028
Qy       5133 GATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGA 5192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 tMetGlyGlyProGlyProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAs 1048
Qy       5193 TGGATCAGGAATAGATACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAA 5252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 pGlySerGlyIleAspThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAs 1068
Qy       5253 CTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACC 5312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 nSerMetGlyProArgMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyPr 1088
Qy       5313 GCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGG 5372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 oProGlyPheAsnProAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerGl 1108
Qy       5373 CTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGAT 5432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 yCysGlySerAlaAsnSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMe 1128
Qy       5433 GGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAG 5492
              ||||||||||||||||||||||||||||||||||||||||
Db       1128 tGlyArgMetProValAsnPheGlySerAsnPheAsnProAsnIle-------------- 1143
Qy       5493 TACCCCAAACACCATACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAA 5552
```

```
Db       1143 ------------------------------------------------------------ 1143
Qy       5553 CAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAA 5612
                                                  ||||||||||||||
Db       1144 ----------------------------------------------GlnArgTyrAlaAs 1148
Qy       5613 CCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGG 5672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1148 nProGlnMetGlyAlaValGlyAsnGlySerProIleCysProProSerAlaSerAspGl 1168
Qy       5673 TACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTC 5732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 yThrProGlyMetProGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSe 1188
Qy       5733 TTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAA 5792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 rSerGlyGluGlnHisGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAs 1208
Qy       5793 CTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTACCCGGCCA 5852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 nPhePheGlnAsnCysAsnGlnMetSerIleValAspGluGluGlyGlyLeuProGlyHi 1228
Qy       5853 TGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCAT 5912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 sAspGlySerMetAsnIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMe 1248
Qy       5913 GCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACA 5972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 tArgProAsnValMetGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGl 1268
Qy       5973 GTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTG 6032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 nSerSerAspGlyIleAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCy 1288
Qy       6033 CAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACA 6092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 sAsnSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGl 1308
Qy       6093 ACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACA 6152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 nHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGl 1328
Qy       6153 AGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTAC 6212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 nGlyGlnIleGlnLeuHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProTh 1348
Qy       6213 TAATAATAATTTAATGTCAACTGCCGGAAGTGTCAGTGCTACTAACGGTGTCTCTGGCAT 6272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 rAsnAsnAsnLeuMetSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIl 1368
Qy       6273 CAATTTCGTAGGTCCCTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCA 6332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 eAsnPheValGlyProSerSerThrAspLeuLysTyrAlaGlnGlnTyrHisSerPheGl 1388
Qy       6333 GCAGCAGTTATATGCTACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCA 6392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 nGlnGlnLeuTyrAlaThrAsnThrArgSerGlnGlnGlnGlnHisMetHisGlnGlnHi 1408
Qy       6393 CCAGAGCAACATGATAACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTCAA 6452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 sGlnSerAsnMetIleThrMetProProAsnLeuSerProAsnProThrPhePheValAs 1428
Qy       6453 CAAA 6456
              ||||
Db       1428 nLys 1429
```

RESULT 2
US-09-915-543-10
; Sequence 10, Application US/09915543
; Publication No. US20020086986A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich

```
;  APPLICANT:  FROESCH, Barbara
;  APPLICANT:  KRAMPS, Thomas
;  APPLICANT:  PETER, Oliver
;  TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
;  TITLE OF INVENTION:  THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
;  FILE REFERENCE: Q60361
;  CURRENT APPLICATION NUMBER: US/09/915,543
;  CURRENT FILING DATE:  2001-07-27
;  PRIOR APPLICATION NUMBER: 60/221,502
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 112
;   TYPE: PRT
;   ORGANISM: Drosophila lgs
US-09-915-543-10

Alignment Scores:
Pred. No.:              2e-38          Length:         112
Score:                  570.00         Matches:        112
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            4.74%          Indels:         0
DB:                     9              Gaps:           0

US-10-664-859-1 (1-6909) x US-09-915-543-10 (1-112)

Qy       4029 TCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCT 4088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnAsnLeuSer 20

Qy       4089 CTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATG 4148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40

Qy       4149 GATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCA 4208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60

Qy       4209 GGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCG 4268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlyThrThrThrValLeuSerAlaAsnLysAsnCysPheGlnAlaAspThrProSerPro 80

Qy       4269 TCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGC 4328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerAsnGlnAsnArgSerArgAsnThrGlySerSerSerValLeuThrHisAsnLeuSer 100

Qy       4329 AGCAACCCAAGTACCCCCTTATCTCATCTATCCCCA 4364
              ||||||||||||||||||||||||||||||||||||
Db        101 SerAsnProSerThrProLeuSerHisLeuSerPro 112

RESULT 3
US-10-322-579-10
; Sequence 10, Application US/10322579
; Publication No. US20030114413A1
; GENERAL INFORMATION:
;  APPLICANT: BASLER, Konrad
;  APPLICANT:  BRUNNER, Erich
;  APPLICANT:  FROESCH, Barbara
;  APPLICANT:  KRAMPS, Thomas
;  APPLICANT:  PETER, Oliver
;  TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
;  TITLE OF INVENTION:  THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
;  FILE REFERENCE: Q60361
;  CURRENT APPLICATION NUMBER: US/10/322,579
;  CURRENT FILING DATE:  2002-12-19
;  PRIOR APPLICATION NUMBER: US/09/915,543
;  PRIOR FILING DATE: 2001-07-27
;  PRIOR APPLICATION NUMBER: 60/221,502
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 112
;   TYPE: PRT
;   ORGANISM: Drosophila lgs
US-10-322-579-10

Alignment Scores:
Pred. No.:              2e-38          Length:         112
Score:                  570.00         Matches:        112
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            4.74%          Indels:         0
DB:                     14             Gaps:           0

US-10-664-859-1 (1-6909) x US-10-322-579-10 (1-112)

Qy       4029 TCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCT 4088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnAsnLeuSer 20

Qy       4089 CTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATG 4148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40

Qy       4149 GATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCA 4208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60

Qy       4209 GGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCG 4268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlyThrThrThrValLeuSerAlaAsnLysAsnCysPheGlnAlaAspThrProSerPro 80

Qy       4269 TCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGC 4328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerAsnGlnAsnArgSerArgAsnThrGlySerSerSerValLeuThrHisAsnLeuSer 100

Qy       4329 AGCAACCCAAGTACCCCCTTATCTCATCTATCCCCA 4364
              ||||||||||||||||||||||||||||||||||||
Db        101 SerAsnProSerThrProLeuSerHisLeuSerPro 112

RESULT 4
US-10-664-859-10
; Sequence 10, Application US/10664859
; Publication No. US20040038901A1
; GENERAL INFORMATION:
;  APPLICANT: BASLER, Konrad
;  APPLICANT:  BRUNNER, Erich
;  APPLICANT:  FROESCH, Barbara
;  APPLICANT:  KRAMPS, Thomas
;  APPLICANT:  PETER, Oliver
;  TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
;  TITLE OF INVENTION:  THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
;  FILE REFERENCE: Q60361
;  CURRENT APPLICATION NUMBER: US/10/664,859
;  CURRENT FILING DATE:  2003-09-22
;  PRIOR APPLICATION NUMBER: US/09/915,543
;  PRIOR FILING DATE: 2001-07-27
;  PRIOR APPLICATION NUMBER: 60/221,502
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 112
;   TYPE: PRT
;   ORGANISM: Drosophila lgs
US-10-664-859-10

Alignment Scores:
Pred. No.:              2e-38          Length:         112
Score:                  570.00         Matches:        112
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            4.74%          Indels:         0
```

```
DB:                    15            Gaps:          0

US-10-664-859-1 (1-6909) x US-10-664-859-10 (1-112)

Qy       4029 TCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCT 4088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnAsnLeuSer 20

Qy       4089 CTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATG 4148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40

Qy       4149 GATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCA 4208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60

Qy       4209 GGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCG 4268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlyThrThrThrValLeuSerAlaAsnLysAsnCysPheGlnAlaAspThrProSerPro 80

Qy       4269 TCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGC 4328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerAsnGlnAsnArgSerArgAsnThrGlySerSerSerValLeuThrHisAsnLeuSer 100

Qy       4329 AGCAACCCAAGTACCCCCTTATCTCATCTATCCCCA 4364
              ||||||||||||||||||||||||||||||||||||
Db        101 SerAsnProSerThrProLeuSerHisLeuSerPro 112

RESULT 5
US-10-276-774-2178
; Sequence 2178, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT:  Tang, Y, Tom et al
;  TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-030
;  CURRENT APPLICATION NUMBER: US/10/276,774
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 2700
;  SOFTWARE: Custom
; SEQ ID NO 2178
;   LENGTH: 1435
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-276-774-2178

Alignment Scores:
Pred. No.:              3.14e-17      Length:        1435
Score:                  321.50        Matches:       313
Percent Similarity:     30.10%        Conservative:  180
Best Local Similarity:  19.11%        Mismatches:    522
Query Match:            2.67%         Indels:        624
DB:                     15            Gaps:          71

US-10-664-859-1 (1-6909) x US-10-276-774-2178 (1-1435)

Qy       2412 TCAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATG 2471
              :::      |||:::   :::   |||         :::                  |||
Db         55 AsnThrGlnSerSerProLysSerLysGlnGluValMetValArgProProThrValMet 74

Qy       2472 AGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCT 2531
              |||      |||      ||||||:::     |||   |||:::
Db         75 SerProSerGlyAsnProGlnLeuAspSerLysPheSerAsnGlnGlyLysGlnGlyGly 94

Qy       2532 GCAATTTCGAGCATATTAGTAAGC---------------------------------- 2555
              :::   |||           |||
Db         95 SerAlaSerGlnSerGlnProSerProCysAspSerLysSerGlyGlyHisThrProLys 114

Qy       2556 ---------GGAGTAGGACCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTA 2606
                      |||   |||         |||:::      |||||||||   |||
Db          115 AlaLeuProGlyProGlyGlySerMetGlyLeuLysAsnGlyAlaGlyAsnGly------ 132

Qy       2607 TTGACTGCCAACGCCAATGGA--------------ATCTCCTCGGGTAGC--------- 2642
                   |||         |||               ||||||:::   |||
Db          133 ------AlaLysGlyLysGlyLysArgGluArgSerIleSerAlaAspSerPheAspGln 150

Qy       2643 ----------------------------------------AGTAATTGTTTGGATTAC 2660
                                                      |||       |||:::
Db          151 ArgAspProGlyThrProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHis 170

Qy       2661 ATGCAACAGCAAAATCAC---------------------------------------- 2678
              ::::::   |||:::
Db          171 IleLysSerGlnAspSerGlnHisThrProHisSerMetThrProSerAsnAlaThrAla 190

Qy       2678 ---------------------------------------------------------- 2678

Db          191 ProArgSerSerThrProProHisGlyGlnThrThrAlaThrGluProThrProAlaGln 210

Qy       2679 -----------------ATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAA 2720
                              ::::::|||||||||||:::::|||||||||   ||||||
Db          211 LysThrProAlaLysValValTyrValPheSerThrGluMetAlaAsnLysAlaAlaGlu 230

Qy       2721 TCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACA 2780
              :::|||||   ||||||   :::||||||:::::::|||
Db          231 AlaValLeuLysGlyGlnValGluThrIleValSerPheHis----------------- 244

Qy       2781 AAAAGCTTCCTGGAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGG 2840
                                     ::::::|||   :::      ||||||   :::|||
Db          245 ----------------------IleGlnAsn---IleSerAsnAsnLysThrGluArg 255

Qy       2841 CACAATTCCGTCGGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAAT 2900
                    :::       |||                                  |||
Db          256 SerThrAla---------ProLeuAsnThrGlnIleSerAlaLeuArgAsnAspProLys 272

Qy       2901 CCTGTAGCCAAAATAACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCC 2960
              |||:::                |||||||||
Db          273 ProLeu------------ProGlnGlnProProAlaProAla----------------- 282

Qy       2961 CAATTCAATCAACATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT 3020
                   ||||||   :::||||||   :::::::|||          |||:::
Db          283 ------AsnGlnAspGlnAsnSerSerGlnAsnThrArgLeuGlnProThrProProIle 300

Qy       3021 ------------------------------------GTCGACCAGTCTGATCCTATG 3041
                                                 :::|||:::      |||
Db          301 ProAlaProAlaProLysProAlaAlaProProArgProLeuAspArgGluSerPro--- 319

Qy       3042 GGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT--------- 3092
              |||   |||   :::|||:::         |||      ::::::::::
Db          320 GlyValGluAsnLysLeuIleProSerValGlySerProAlaSerSerThrProLeuPro 339

Qy       3093 ------AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAG 3146
                    :::|||   ||||||      :::      :::::::::      |||:::
Db          340 ProAspGlyThrGlyProAsnSerThrProAsnAsnArgAlaValThrProValSerGln 359

Qy       3147 TCACAGGCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGA 3206
                                  :::                        |||:::   |||
Db          360 GlySerAsnSerSerSerAlaAspProLysAlaProProProProProValSerSerGly 379

Qy       3207 AGCGATCCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCA 3266
                   |||   |||      ||||||                                 |||
Db          380 GluProProThrLeu---GlyGluAsn----------------------------Pro 388

Qy       3267 GACGAAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAA 3326
              |||      |||:::   :::|||   :::|||||||||      |||      ::::::
Db          389 Asp---GlyLeuSerGlnGluGlnLeuGluHisArgGluArgSerLeuGlnThrLeuArg 407

Qy       3327 AAAATGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCA 3386
                 :::   :::   |||||||||:::   :::      ||||||
Db          408 AspIleGlnArgMetLeuPheProAspGluLysGluPheThrGlyAlaGln--------- 424

Qy       3387 CAGATAACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCT 3446
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 21:47:47 ; Search time 164 Seconds
(without alignments)
6289.639 Million cell updates/sec

Title: US-10-664-859-1
Perfect score: 12037
Sequence: 1 acgagtgcttctcttattat.....aaaaaaaaaaaaaaaaaaa 6909

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10664859/runat_02112005_171804_3553/app_query.fasta_1.7047
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10664859_@CGN_1_1_181_@runat_02112005_171804_3553 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2429.5	20.2	506	4	US-09-270-767-44797	Sequence 44797, A
2	1573	13.1	317	4	US-09-270-767-60260	Sequence 60260, A
3	1145	9.5	259	4	US-09-270-767-46113	Sequence 46113, A
4	935	7.8	175	4	US-09-270-767-36663	Sequence 36663, A
5	935	7.8	175	4	US-09-270-767-51880	Sequence 51880, A
6	599.5	5.0	230	4	US-09-270-767-44999	Sequence 44999, A
7	594	4.9	140	4	US-09-270-767-36764	Sequence 36764, A
8	594	4.9	140	4	US-09-270-767-51981	Sequence 51981, A
c 9	312	2.6	65	4	US-09-270-767-43686	Sequence 43686, A
10	238.5	2.0	1179	4	US-09-949-016-7088	Sequence 7088, Ap
11	236	2.0	43	4	US-09-270-767-61668	Sequence 61668, A
12	227	1.9	1076	4	US-09-949-016-7421	Sequence 7421, Ap
13	222	1.8	1420	4	US-09-125-635-4	Sequence 4, Appli
14	219	1.8	2451	4	US-09-949-016-9675	Sequence 9675, Ap
15	216.5	1.8	1061	3	US-08-701-154A-5	Sequence 5, Appli
16	204	1.7	1584	3	US-09-457-040B-27	Sequence 27, Appl
17	203	1.7	2314	3	US-09-268-347-49	Sequence 49, Appl
18	202	1.7	1600	2	US-08-617-697-10	Sequence 10, Appl
19	201	1.7	1036	3	US-08-891-640-3	Sequence 3, Appli
20	201	1.7	1302	4	US-09-949-016-10852	Sequence 10852, A
21	198.5	1.6	1258	4	US-09-538-092-886	Sequence 886, App
22	198.5	1.6	1531	4	US-09-949-016-6740	Sequence 6740, Ap
23	198	1.6	2703	1	US-08-185-432-19	Sequence 19, Appl
24	198	1.6	2703	4	US-08-899-232-4	Sequence 4, Appli
25	198	1.6	2703	4	US-09-121-457-4	Sequence 4, Appli
26	196.5	1.6	2504	4	US-09-328-352-5821	Sequence 5821, Ap
27	194.5	1.6	1529	2	US-08-728-470-10	Sequence 10, Appl
28	194.5	1.6	1529	3	US-08-719-641-10	Sequence 10, Appl
29	193	1.6	908	4	US-09-949-016-8542	Sequence 8542, Ap
30	193	1.6	2169	4	US-09-949-016-6930	Sequence 6930, Ap
31	192	1.6	2736	4	US-09-252-991A-30227	Sequence 30227, A
32	191	1.6	3969	3	US-08-061-376-5	Sequence 5, Appli
33	191	1.6	3969	4	US-09-538-092-1262	Sequence 1262, Ap
34	189	1.6	1268	4	US-09-270-767-45320	Sequence 45320, A
35	186.5	1.5	1776	4	US-09-556-877-179	Sequence 179, App
36	186.5	1.5	1776	4	US-09-620-412C-179	Sequence 179, App
37	186.5	1.5	1776	4	US-09-598-419-179	Sequence 179, App
38	183.5	1.5	1187	4	US-09-949-016-6513	Sequence 6513, Ap
39	183.5	1.5	1402	4	US-09-445-353E-2	Sequence 2, Appli
40	181.5	1.5	1477	4	US-09-206-942-71	Sequence 71, Appl
41	180.5	1.5	1992	4	US-08-621-944A-3	Sequence 3, Appli
42	180.5	1.5	1992	4	US-08-945-567D-3	Sequence 3, Appli
43	179.5	1.5	1833	4	US-08-621-944A-4	Sequence 4, Appli
44	179.5	1.5	1833	4	US-08-945-567D-4	Sequence 4, Appli
45	176.5	1.5	1323	4	US-09-248-796A-19543	Sequence 19543, A

ALIGNMENTS

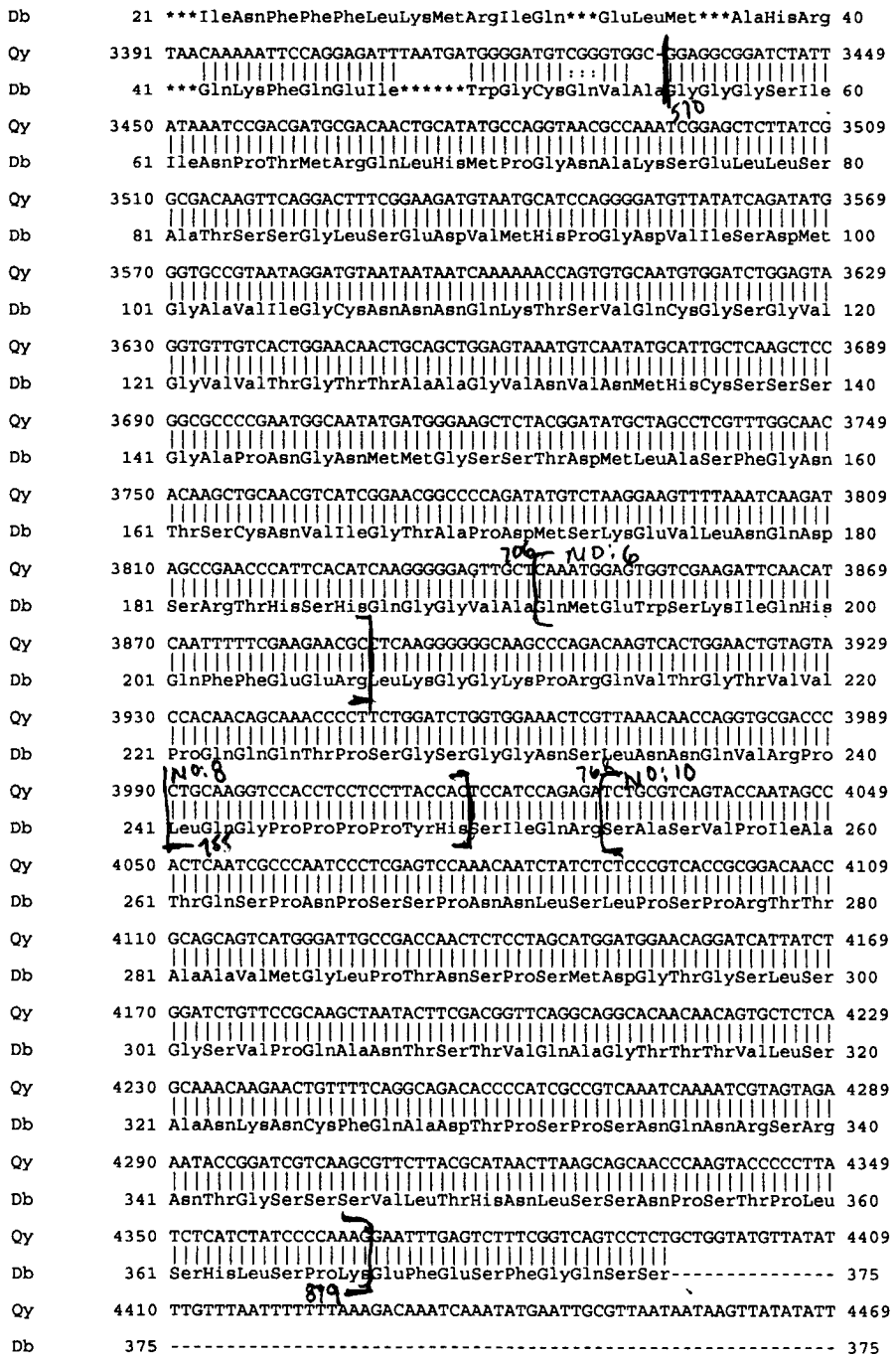
RESULT 1
US-09-270-767-44797
; Sequence 44797, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44797
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44797

Alignment Scores:

Pred. No.:	6.03e-220	Length:	506
Score:	2429.50	Matches:	495
Percent Similarity:	80.03%	Conservative:	2
Best Local Similarity:	79.71%	Mismatches:	9
Query Match:	20.18%	Indels:	116
DB:	4	Gaps:	2

US-10-664-859-1 (1-6909) x US-09-270-767-44797 (1-506)

```
Qy       3271 AAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAA 3330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LysThrLeuHisHisAsnSerGlyAsnIleGlyLysAsnSerTrpGlnLys***LysLys 20

Qy       3331 TGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGA 3390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



RESULT 2
US-09-270-767-60260
; Sequence 60260, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60260
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-60260

Alignment Scores:

Pred. No.:	2.8e-139	Length:	317
Score:	1573.00	Matches:	308
Percent Similarity:	97.48%	Conservative:	1
Best Local Similarity:	97.16%	Mismatches:	8
Query Match:	13.07%	Indels:	1
DB:	4	Gaps:	0

US-10-664-859-1 (1-6909) x US-09-270-767-60260 (1-317)

```
Qy       3271 AAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAA 3330
Db          1 LysThrLeuHisHisAsnSerGlyAsnIleGlyLysAsnSerTrpGlnLys***LysLys 20

Qy       3331 TGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGA 3390
Db         21 ***IleAsnPhePhePheLeuLysMetArgIleGln***GluLeuMet***AlaHisArg 40

Qy       3391 TAACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGC-GGAGGCGGATCTATT 3449
Db         41 ***GlnLysPheGlnGluIle******TrpGlyCysGlnValAlaGlyGlyGlySerIle 60

Qy       3450 ATAAATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCG 3509
Db         61 IleAsnProThrMetArgGlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSer 80

Qy       3510 GCGACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATG 3569
Db         81 AlaThrSerSerGlyLeuSerGluAspValMetHisProGlyAspValIleSerAspMet 100

Qy       3570 GGTGCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTA 3629
Db        101 GlyAlaValIleGlyCysAsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyVal 120

Qy       3630 GGTGTTGTCACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCC 3689
Db        121 GlyValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSer 140

Qy       3690 GGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAAC 3749
Db        141 GlyAlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsn 160

Qy       3750 ACAAGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGAT 3809
Db        161 ThrSerCysAsnValIleGlyThrAlaProAspMetSerLysGluValLeuAsnGlnAsp 180

Qy       3810 AGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACAT 3869
Db        181 SerArgThrHisSerHisGlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHis 200

Qy       3870 CAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTA 3929
Db        201 GlnPhePheGluGluArgLeuLysGlyGlyLysProArgGlnValThrGlyThrValVal 220

Qy       3930 CCACAACAGCAAACCCCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCC 3989
Db        221 ProGlnGlnGlnThrProSerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgPro 240

Qy       3990 CTGCAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCC 4049
Db        241 LeuGlnGlyProProProProTyrHisSerIleGlnArgSerAlaSerValProIleAla 260

Qy       4050 ACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACC 4109
Db        261 ThrGlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThr 280

Qy       4110 GCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCT 4169
Db        281 AlaAlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSer 300

Qy       4170 GGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGCACAACAACA 4220
Db        301 GlySerValProGlnAlaAsnThrSerThrValGlnAlaGlyThrThrThr 317
```

RESULT 3
US-09-270-767-46113
; Sequence 46113, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46113
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46113

Alignment Scores:

Pred. No.:	5.83e-99	Length:	259
Score:	1145.00	Matches:	250
Percent Similarity:	50.60%	Conservative:	1
Best Local Similarity:	50.40%	Mismatches:	1
Query Match:	9.51%	Indels:	244
DB:	4	Gaps:	1

US-10-664-859-1 (1-6909) x US-09-270-767-46113 (1-259)

```
Qy       1453 TCAGAAATAAGTTCTAATAAAGCAAAAGGTCAAGCAGCTGGTGGCGGCTGCGAAACAGGT 1512
Db          8 AlaGluIleSerSerAsnLysAlaLysGlyGlnAlaAlaGlyGlyGlyCysGluThrGly 27

Qy       1513 TCTACATCCAGTTTGACTGTCAAGGAAGAACCCACCGATGTCTTAGGCAGTTTAGTAAAT 1572
Db         28 SerThrSerSerLeuThrValLysGluGluProThrAspValLeuGlySerLeuValAsn 47

Qy       1573 ATGAAAAAAGAAGAAAGAGAAAATCATTCGCCAACGATGTCCCCTGTTGGTTTTGGTTCA 1632
Db         48 MetLysLysGluGluArgGluAsnHisSerProThrMetSerProValGlyPheGlySer 67

Qy       1633 ATTGGTAATGCACAGGACAACTCCGCTACACCGGGTAAGTTTTAAGAGATCCATATAAAG 1692
Db         68 IleGlyAsnAlaGlnAspAsnSerAlaThr---------------------------- 77

Qy       1693 CAAATAACAAGAATTAATGTCAGTTACCAATTTTATTTGATAGTCAAAGAACTACTATAG 1752
Db         77 ---------------------------------------------------------- 77

Qy       1753 CGATATCTCCTGCCTTTTAATTTTATTTTAATTAGGAAATACGAATATTTCTAATTTGTA 1812
Db         77 ---------------------------------------------------------- 77

Qy       1813 AAATAAAATTGATTAATTAACTAGAATTTAAAAACCTTTTGAATTAGGACATACCCTTCC 1872
Db         77 ---------------------------------------------------------- 77

Qy       1873 AAAAATCAGTAATCATTGGGAACGAGAGTGTGGTCCCGAAGGAGACTACTATAAAACCTT 1932
Db         77 ---------------------------------------------------------- 77

Qy       1933 TTGAGCTATCTGATACTGCACGCTACTAAAAATGATTAGTTTAGGAAAATGGGTGTAATT 1992
Db         77 ---------------------------------------------------------- 77

Qy       1993 TTGTAGGAAGTTTTCATTTTAGAAGAAATGTGATTATTTTATTAAACCCCTTCAAGCGGA 2052
Db         77 ---------------------------------------------------------- 77

Qy       2053 ACTACATTTGTTCTACGATATTTTGGAAAAACAAATGGTTAAGTTGGAAAGTGCCTATAA 2112
Db         77 ---------------------------------------------------------- 77

Qy       2113 AACAGAATTCCACGGTTTCAAATACTAACCAGGTTTTTGATTTAATTTTGATTAAATGAG 2172
Db         77 ---------------------------------------------------------- 77

Qy       2173 AAATTATCACACTTCAGTTAAAATGTTTAATTCGATTAAGGTCGGACAATCACAGCAGAT 2232
Db         77 ---------------------------------------------------------- 77
```

```
Qy       2233 TTCCATTTTTGCGTGTATATATAGAAGTCGCCTTCACACTCTTCTGGCGCGCTTCACCAC 2292

Db         77 ------------------------------------------------------------ 77

Qy       2293 TACGTGGAGTTCCGCCCGCAGTGATTTATATAGATGATTTACGAGTTATTTAATTTTTTA 2352

Db         77 ------------------------------------------------------------ 77

Qy       2353 TGGTGTATTTTAATAAATATCTTATTTATTCATTTTACATAGTTAAAATTGAAAGAATTT 2412
                                                  ||||||||||||||||||
Db         78 ----------------------------------Pro--ValLysIleGluArgIleS  85

Qy       2413 CAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATGA 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 erAsnAspSerThrThrGluLysLysGlySerSerLeuThrMetAsnAsnAspGluMetS 105

Qy       2473 GCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTG 2532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 erMetGluGlyCysAsnGlnLeuAsnProAspPheIleAsnGluSerLeuAsnAsnProA 125

Qy       2533 CAATTTCGAGCATATTAGTAAGCGGAGTAGGACCAATACCCGGAATCGGAGTTGGAGCGG 2592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 laIleSerSerIleLeuValSerGlyValGlyProIleProGlyIleGlyValGlyAlaG 145

Qy       2593 GGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCCTCGGGTAGCAGTAATTGTT 2652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 lyThrGlyAsnLeuLeuThrAlaAsnAlaAsnGlyIleSerSerGlySerSerAsnCysL 165

Qy       2653 TGGATTACATGCAACAGCAAAATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAG 2712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 euAspTyrMetGlnGlnGlnAsnHisIlePheValPheSerThrGlnLeuAlaAsnLysG 185

Qy       2713 GGGCCGAATCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGC 2772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 lyAlaGluSerValLeuSerGlyGlnPheGlnThrIleIleAlaTyrHisCysThrGlnP 205

Qy       2773 CTGCTACAAAAAGCTTCCTGGAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGT 2832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 roAlaThrLysSerPheLeuGluAspPhePheMetLysAsnProLeuLysIleAsnLysL 225

Qy       2833 TACAGCGGCACAATTCCGTCGGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTC 2892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 euGlnArgHisAsnSerValGlyMetProTrpIleGlyMetGlyGlnValGlyLeuThrP 245

Qy       2893 CTCCTAATCCTGTAGCCAAAATAACACAACAGCAGCCACATACA 2936
              |||||||||||||||||||||||||||||||||||||||||||
Db        245 roProAsnProValAlaLysIleThrGlnGlnGlnProHisThr 259
```

RESULT 4
US-09-270-767-36663
; Sequence 36663, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36663
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-36663

Alignment Scores:

Pred. No.:	2.97e-79	Length:	175
Score:	935.00	Matches:	175
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	7.77%	Indels:	0
DB:	4	Gaps:	0

US-10-664-859-1 (1-6909) x US-09-270-767-36663 (1-175)

```
Qy       5725 ATGAATTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAAT 5784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetAsnSerSerGlyGluGlnHisGlnAsnLysIleThrAsnAsnProGlyAlaSerAsn 20

Qy       5785 GGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTA 5844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyIleAsnPhePheGlnAsnCysAsnGlnMetSerIleValAspGluGluGlyGlyLeu 40

Qy       5845 CCCGGCCATGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA 5904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProGlyHisAspGlySerMetAsnIleGlyGlnProSerMetIleArgGlyMetArgPro 60

Qy       5905 CATGCCATGCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAG 5964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HisAlaMetArgProAsnValMetGlyAlaArgMetProProValAsnArgGlnIleGln 80

Qy       5965 TTTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAAC 6024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 PheAlaGlnSerSerAspGlyIleAspCysValGlyAspProSerSerPhePheThrAsn 100

Qy       6025 GCTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCT 6084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AlaSerCysAsnSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAsnGlnPro 120

Qy       6085 AAGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCA 6144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LysThrGlnHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAla 140

Qy       6145 GTGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATT 6204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ValAlaGlnGlyGlnIleGlnLeuHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIle 160

Qy       6205 GGACCTACTAATAATAATTTAATGTCAACTGCCGGAAGTGTCAGT 6249
              |||||||||||||||||||||||||||||||||||||||||||||
Db        161 GlyProThrAsnAsnAsnLeuMetSerThrAlaGlySerValSer 175
```

RESULT 5
US-09-270-767-51880
; Sequence 51880, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51880
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-51880

Alignment Scores:

Pred. No.:	2.97e-79	Length:	175
Score:	935.00	Matches:	175
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	7.77%	Indels:	0
DB:	4	Gaps:	0

US-10-664-859-1 (1-6909) x US-09-270-767-51880 (1-175)

```
Qy       5725 ATGAATTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAAT 5784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetAsnSerSerGlyGluGlnHisGlnAsnLysIleThrAsnAsnProGlyAlaSerAsn 20

Qy       5785 GGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTA 5844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         21 GlyIleAsnPhePheGlnAsnCysAsnGlnMetSerIleValAspGluGluGlyGlyLeu 40

Qy       5845 CCCGGCCATGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA 5904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ProGlyHisAspGlySerMetAsnIleGlyGlnProSerMetIleArgGlyMetArgPro 60

Qy       5905 CATGCCATGCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAG 5964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HisAlaMetArgProAsnValMetGlyAlaArgMetProProValAsnArgGlnIleGln 80

Qy       5965 TTTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAAC 6024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 PheAlaGlnSerSerAspGlyIleAspCysValGlyAspProSerSerPhePheThrAsn 100

Qy       6025 GCTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCT 6084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 AlaSerCysAsnSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAsnGlnPro 120

Qy       6085 AAGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCA 6144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LysThrGlnHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAla 140

Qy       6145 GTGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATT 6204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 ValAlaGlnGlyGlnIleGlnLeuHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIle 160

Qy       6205 GGACCTACTAATAATAATTTAATGTCAACTGCCGGAAGTGTCAGT 6249
              |||||||||||||||||||||||||||||||||||||||||||||
Db        161 GlyProThrAsnAsnAsnLeuMetSerThrAlaGlySerValSer 175
```

RESULT 6
US-09-270-767-44999
; Sequence 44999, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44999
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44999

Alignment Scores:

Pred. No.:	1.54e-47	Length:	230
Score:	599.50	Matches:	118
Percent Similarity:	93.13%	Conservative:	4
Best Local Similarity:	90.08%	Mismatches:	6
Query Match:	4.98%	Indels:	3
DB:	4	Gaps:	1

US-10-664-859-1 (1-6909) x US-09-270-767-44999 (1-230)

```
Qy       2382 TCATTTTACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAAAAAAAAGGA 2441
              |||||||||         ::::::   :::|||   :::              |||||||||
Db        103 SerPheTyr***SerThrValLysLeuMetSer***AsnGly---------LysLysGly 119

Qy       2442 TCGTCCTTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCC 2501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 SerSerLeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGlnLeuAsnPro 139

Qy       2502 GATTTTATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTA 2561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 AspPheIleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuValSerGlyVal 159

Qy       2562 GGACCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCC 2621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GlyProIleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThrAlaAsnAla 179

Qy       2622 AATGGAATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCACATA 2681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 AsnGlyIleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGlnAsnHisIle 199

Qy       2682 TTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTT 2741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 PheValPheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSerGlyGlnPhe 219

Qy       2742 CAAACTATTATTGCGTATCACTGCACTCAGCCT 2774
              |||||||||||||||||||||||||||||||||
Db        220 GlnThrIleIleAlaTyrHisCysThrGlnPro 230
```

RESULT 7
US-09-270-767-36764
; Sequence 36764, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36764
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-36764

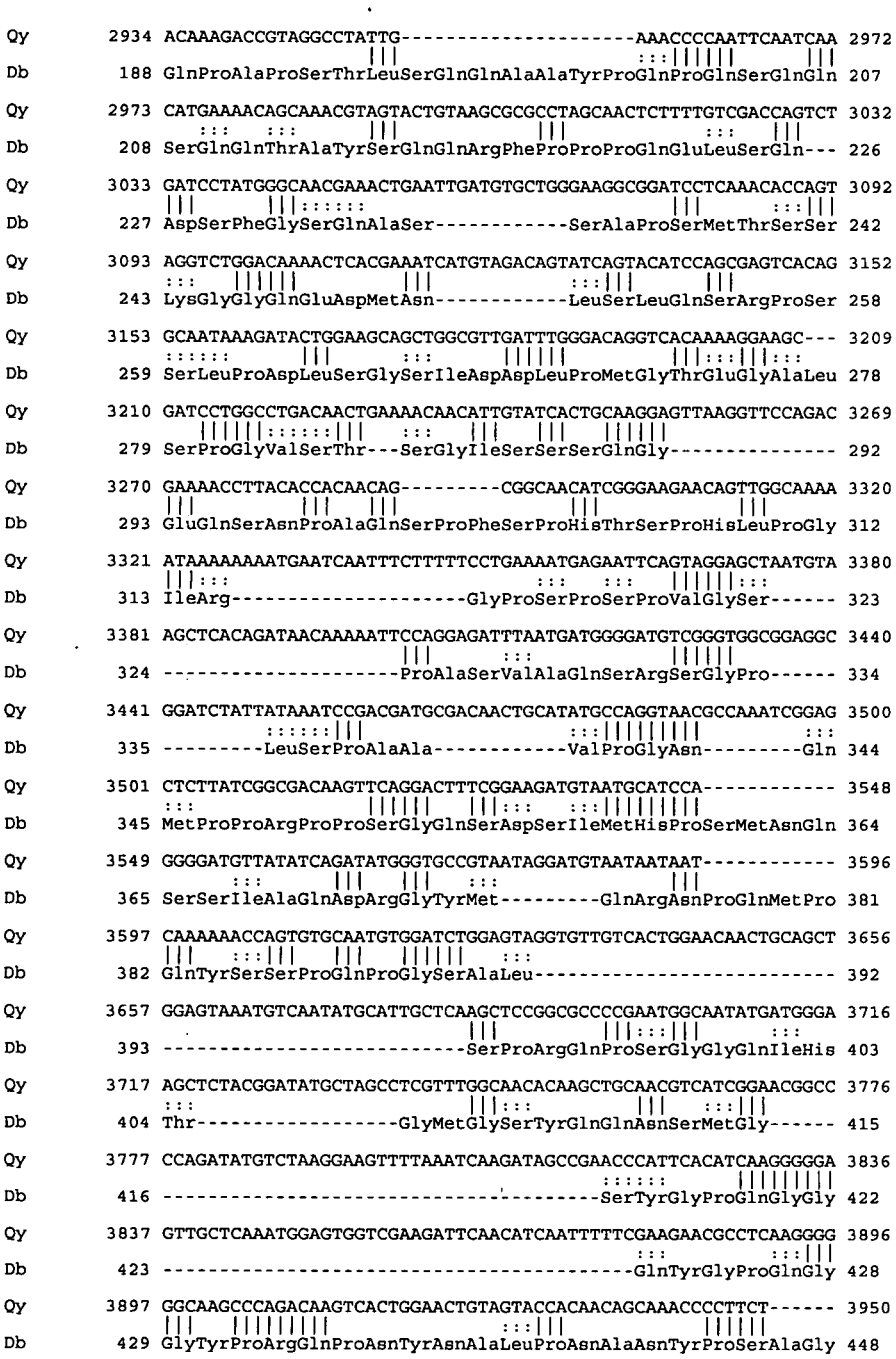
Alignment Scores:

Pred. No.:	3.8e-47	Length:	140
Score:	594.00	Matches:	123
Percent Similarity:	88.57%	Conservative:	1
Best Local Similarity:	87.86%	Mismatches:	16
Query Match:	4.93%	Indels:	0
DB:	4	Gaps:	0

US-10-664-859-1 (1-6909) x US-09-270-767-36764 (1-140)

```
Qy       6391 CACCAGAGCAACATGATAACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTC 6450
              |||||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db          1 HisGlnSerAsnMetIleThrMetProSerAsnLeuSerProAsnProThrPhePheVal 20

Qy       6451 AACAAATAAACTTCTAAATTTTTGCCGCCCTCGTCATGTATTGTTTACTAGTCTCCAAAT 6510
              ||||||   ||||||||||||||||||          ||||||||||||   |||||||||
Db         21 AsnLys***ThrSerLysPheLeuProSerLeuProCysIleValTyr***SerProAsn 40

Qy       6511 TAAGACATGCATCTCTAAATAAGATTTTTTGAAGCTTATTTACTTAGGTGTTTTTACAAC 6570
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         41 ***AspMetHisLeu***IleArgPhePheGluAlaTyrLeuLeuArgCysPheTyrAsn 60

Qy       6571 GGAGAAAATAAACTTTTGGATATGCAAATGATAACGTTGGAAACAACATAATTCATTTGC 6630
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         61 GlyGluAsnLysLeuLeuAspMetGlnMetIleThrLeuGluThrThr***PheIleCys 80

Qy       6631 AACTTTTAGAAGTCACGTCGAAGTTAAATGTAGAATCTGTATTTTAACATAATAGGTCAT 6690
              ||||||   ||||||||||||||   |||   ||||||||||||||||||||||||||||
Db         81 AsnPhe***LysSerArgArgSer***Met***AsnLeuTyrPheAsnIleIleGlyHis 100

Qy       6691 CTGTAAAAATAATTAAACATCGAAATTTTAGTTATCAGCAGCTATTTTCTGTTATTATTT 6750
              |||   |||   |||||||||||||||||||||||||||||:::||||||||||||||
Db        101 LeuLeuLys***LeuAsnIleGluIleLeuValIleSerSerTyrTyrLeuLeuLeuPhe 120

Qy       6751 AATATGTGCGCTGCTCTCTCTGTGTTAAATGAAATTAAAATATATATATAAATGTAAAAC 6810
              ||||||||||||||||||||||||||||||||||||||||||||||||||   |||   |||
Db        121 AsnMetCysAlaAlaLeuSerValLeuAsnGluIleLysIleTyrIle***Met***Asn 140
```

```
RESULT 8
US-09-270-767-51981
; Sequence 51981, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51981
;   LENGTH: 140
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-51981

Alignment Scores:
Pred. No.:              3.8e-47       Length:         140
Score:                  594.00        Matches:        123
Percent Similarity:     88.57%        Conservative:   1
Best Local Similarity:  87.86%        Mismatches:     16
Query Match:            4.93%         Indels:         0
DB:                     4             Gaps:           0

US-10-664-859-1 (1-6909) x US-09-270-767-51981 (1-140)

Qy       6391 CACCAGAGCAACATGATAACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTC 6450
              ||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db          1 HisGlnSerAsnMetIleThrMetProSerAsnLeuSerProAsnProThrPhePheVal 20

Qy       6451 AACAAATAAACTTCTAAATTTTTGCCGCCCTCGTCATGTATTGTTTACTAGTCTCCAAAT 6510
              ||||||   ||||||||||||||||||         ||||||||||||   |||||||||
Db         21 AsnLys***ThrSerLysPheLeuProSerLeuProCysIleValTyr***SerProAsn 40

Qy       6511 TAAGACATGCATCTCTAAATAAGATTTTTTGAAGCTTATTTACTTAGGTGTTTTTACAAC 6570
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         41 ***AspMetHisLeu***IleArgPhePheGluAlaTyrLeuLeuArgCysPheTyrAsn 60

Qy       6571 GGAGAAAATAAACTTTTGGATATGCAAATGATAACGTTGGAAACAACATAATTCATTTGC 6630
              ||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         61 GlyGluAsnLysLeuLeuAspMetGlnMetIleThrLeuGluThrThr***PheIleCys 80

Qy       6631 AACTTTTAGAAGTCACGTCGAAGTTAAATGTAGAATCTGTATTTTAACATAATAGGTCAT 6690
              ||||||   ||||||||||||||   |||   ||||||||||||||||||||||||||
Db         81 AsnPhe***LysSerArgArgSer***Met***AsnLeuTyrPheAsnIleIleGlyHis 100

Qy       6691 CTGTAAAAATAATTAAACATCGAAATTTTAGTTATCAGCAGCTATTTTCTGTTATTATTT 6750
              |||   |||   ||||||||||||||||||||||||||||:::||||||||||||||
Db        101 LeuLeuLys***LeuAsnIleGluIleLeuValIleSerSerTyrTyrLeuLeuLeuPhe 120

Qy       6751 AATATGTGCGCTGCTCTCTCTGTGTTAAATGAAATTAAAATATATATATAAATGTAAAAC 6810
              ||||||||||||||||||||||||||||||||||||||||||||||   |||   |||
Db        121 AsnMetCysAlaAlaLeuSerValLeuAsnGluIleLysIleTyrIle***Met***Asn 140

RESULT 9
US-09-270-767-43686
; Sequence 43686, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43686
;   LENGTH: 65
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43686

Alignment Scores:
Pred. No.:              9.59e-21      Length:         65
Score:                  312.00        Matches:        63
Percent Similarity:     96.92%        Conservative:   0
Best Local Similarity:  96.92%        Mismatches:     2
Query Match:            2.58%         Indels:         1
DB:                     4             Gaps:           0

US-10-664-859-1 (1-6909) x US-09-270-767-43686 (1-65)

Qy       4766 ATCGTACATCTTTATTTGCCTCTATTAGACTATTTACTGGTGACCGCTGACCCTGTGGGC 4707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IleValHisLeuTyrLeuProLeuLeuAspTyrLeuLeuValThrAlaAspProValGly 20

Qy       4706 TTGGTCGCCTACTTTTCATGTTATCACCTGTAGAAATATTATAATAAACATATTTTTAAT 4647
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db         21 LeuValAlaTyrPheSerCysTyrHisLeu***LysTyrTyrAsnLysHisIlePheAsn 40

Qy       4646 TTCTTATTAAACAATAGTGATTTTCAATTAAC-GGATATGAATTTTCGTGTTTAATTTTA 4588
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db         41 PheLeuLeuAsnAsnSerAspPheGlnLeuThrGlyTyrGluPheSerCysLeuIleLeu 60

Qy       4587 ATAGTAGATTATCTC 4573
              |||||||||||||||
Db         61 IleValAspTyrLeu 65

RESULT 10
US-09-949-016-7088
; Sequence 7088, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7088
;   LENGTH: 1179
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7088

Alignment Scores:
Pred. No.:              4.52e-13      Length:         1179
Score:                  238.50        Matches:        250
Percent Similarity:     30.38%        Conservative:   137
Best Local Similarity:  19.62%        Mismatches:     406
Query Match:            1.98%         Indels:         481
DB:                     4             Gaps:           64

US-10-664-859-1 (1-6909) x US-09-949-016-7088 (1-1179)

Qy       2814 CCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATAGGCATG 2873
              |||   :::   :::     |||:::|||               |||:::
Db        156 ProSerGlnGlnSerThrThrGlnGlnHisProGlnSerGlnProProTyr--------- 172

Qy       2874 GGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAGCCACAT 2933
                           :::   |||         :::       ||||||||||||
Db        173 --------------SerGlnProGlnAlaGlnSerProTyrGlnGlnGlnGlnProGln 187
```


Qy 2934 ACAAAGACCGTAGGCCTATTG----------AAACCCCAATTCAATCAA 2972
Db 188 GlnProAlaProSerThrLeuSerGlnGlnAlaAlaTyrProGlnProGlnSerGlnGln 207
Qy 2973 CATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCT 3032
Db 208 SerGlnGlnThrAlaTyrSerGlnGlnArgPheProProProGlnGluLeuSerGln--- 226
Qy 3033 GATCCTATGGGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT 3092
Db 227 AspSerPheGlySerGlnAlaSer----------SerAlaProSerMetThrSerSer 242
Qy 3093 AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAG 3152
Db 243 LysGlyGlyGlnGluAspMetAsn----------LeuSerLeuGlnSerArgProSer 258
Qy 3153 GCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGC--- 3209
Db 259 SerLeuProAspLeuSerGlySerIleAspAspLeuProMetGlyThrGluGlyAlaLeu 278
Qy 3210 GATCCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGAC 3269
Db 279 SerProGlyValSerThr---SerGlyIleSerSerSerGlnGly---------- 292
Qy 3270 GAAAACCTTACACCACAACAG---------CGGCAACATCGGGAAGAACAGTTGGCAAAA 3320
Db 293 GluGlnSerAsnProAlaGlnSerProPheSerProHisThrSerProHisLeuProGly 312
Qy 3321 ATAAAAAAAAATGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTA 3380
Db 313 IleArg----------GlyProSerProSerProValGlySer------ 323
Qy 3381 AGCTCACAGATAACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGC 3440
Db 324 ----------ProAlaSerValAlaGlnSerArgSerGlyPro------ 334
Qy 3441 GGATCTATTATAAATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAG 3500
Db 335 ---------LeuSerProAlaAla----------ValProGlyAsn---------Gln 344
Qy 3501 CTCTTATCGGCGACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCA---------- 3548
Db 345 MetProProArgProProSerGlyGlnSerAspSerIleMetHisProSerMetAsnGln 364
Qy 3549 GGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGTAATAATAAT---------- 3596
Db 365 SerSerIleAlaGlnAspArgGlyTyrMet---------GlnArgAsnProGlnMetPro 381
Qy 3597 CAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACAACTGCAGCT 3656
Db 382 GlnTyrSerSerProGlnProGlySerAlaLeu---------- 392
Qy 3657 GGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAATATGATGGGA 3716
Db 393 ----------SerProArgGlnProSerGlyGlyGlnIleHis 403
Qy 3717 AGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATCGGAACGGCC 3776
Db 404 Thr----------GlyMetGlySerTyrGlnGlnAsnSerMetGly------ 415
Qy 3777 CCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACCCATTCACATCAAGGGGGA 3836
Db 416 ----------SerTyrGlyProGlnGlyGly 422
Qy 3837 GTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAATTTTTCGAAGAACGCCTCAAGGGG 3896
Db 423 ----------GlnTyrGlyProGlnGly 428
Qy 3897 GGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACCCCTTCT------ 3950
Db 429 GlyTyrProArgGlnProAsnTyrAsnAlaLeuProAsnAlaAsnTyrProSerAlaGly 448
Qy 3951 ----------GGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGA 3986
Db 449 MetAlaGlyGlyIleAsnProMetGlyAlaGlyGlyGln---------- 461
Qy 3987 CCCCTGCAAGGTCCACCT------CCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTA 4040
Db 462 ---MetHisGlyGlnProGlyIleProProTyrGlyThrLeuPro---ProGlyArgMet 479
Qy 4041 CCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTCCCG------ 4094
Db 480 SerHisAlaSerMetGlyAsnArgProTyrGlyProAsnMetAlaAsnMetProProGln 499
Qy 4095 ----------TCACCGCGGACAACCGCA 4112
Db 500 ValGlySerGlyMetCysProProProGlyGlyMetAsnArgLysThrGlnGluThrAla 519
Qy 4113 GCAGTCATGGGATTGCCGACCAACTCT----------CCTAGCATG 4148
Db 520 ValAlaMetHisValAlaAlaAsnSerIleGlnAsnArgProProGlyTyrProAsnMet 539
Qy 4149 GAT----------GGAACAGGA----------TCATTATCT 4169
Db 540 AsnGlnGlyGlyMetMetGlyThrGlyProProTyrGlyGlnGlyIleAsnSerMetAla 559
Qy 4170 GGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCA 4229
Db 560 GlyMetIleAsnProGlnGlyProProTyrSerMetGlyGlyThr----------Met 575
Qy 4230 GCAAACAAGAACTGTTTTCAGGCAGACACCCCA---------- 4262
Db 576 AlaAsnAsnSerAlaGlyMetAlaAlaSerProGluMetMetGlyLeuGlyAspValLys 595
Qy 4263 ---TCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCAT 4319
Db 596 LeuThrProAlaThrLysMetAsnAsnLysAlaAspGlyThrProLysThrGluSerLys 615
Qy 4320 AACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAGGAATTTGAGTCT 4379
Db 616 SerLysLysSerSerSerSerThrThrThrAsnGluLysIleThrLysLeuTyrGluLeu 635
Qy 4380 TTCGGTCAGTCCTCTGCTGGTATG---------TTATATTTGTTTAATTTTTTTAAAGAC 4430
Db 636 GlyGlyGluProGluArgLysMetTrpValAspArgTyrLeu---AlaPheThrGluGlu 654
Qy 4431 AAA----------TCAAATATGAATTGCGTTAATAATAAG---------- 4460
Db 655 LysAlaMetGlyMetThrAsnLeuProAlaValGlyArgLysProLeuAspLeuTyrArg 674
Qy 4461 TTATATATTACA------TAACTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAAT 4514
Db 675 LeuTyrValSerValLysGluIleGlyGlyLeuThrGlnValAsnLysAsnLysLysTrp 694
Qy 4515 AAATTATTTTCCGGACCGCCCATCCATTTCTTGAATCCAATTTCTGGAGTGATTGTTAGA 4574
Db 695 Arg---------- 695
Qy 4575 GATAATCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAATCACTATT 4634
Db 696 ---GluLeuAlaThrAsnLeuAsnValGlyThrSer-SerSerAlaAlaSerSerLeu-- 713
Qy 4635 GTTTAATAAGAAATTAAAAATATGTTTATTATAAT---------- 4669
Db 714 ----LysLysGlnTyrIleGlnCysLeu-TyrAlaPheGluCysLysIleGluArgGlyG 732
Qy 4670 ----------ATTTCTACAGGTGATAACATGAAAAGT----------A 4697
Db 732 luAspProProProAspIlePheAlaAlaAlaAspSerLysLysSerGlnProLysIleG 752
Qy 4698 GGCGACCAAGCCCA----------CAGGGTCAGCGGTCACCAGTAAATAGTC 4739
Db 752 lnProProSerProAlaGlySerGlySerMetGlnGlyProGlnThrProGlnSerThrS 772
Qy 4740 TAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTTTTAACCCGCAT- 4798

```
Db        772 erSerSerMetAlaGluGlyGlyAspLeuLysProProThrProAlaSerThrProHisS 792

Qy       4799 ----------------CCACATATG---CAAAGCAATTCAAATTCAGCATTAAACGCCT 4838

Db        792 erGlnIleProProLeuProGlyMetSerArgSerAsnSerValGlyIleGlnAspAlaP 812

Qy       4839 ATAAAATGGGCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGT 4898

Db        812 heAsnAspGlySerAspSer--Thr---------------------------------- 819

Qy       4899 TTTTGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACA 4958

Db        820 ----------------------PheGlnLysArgAsnSer------------------- 825

Qy       4959 ATTTAGTCGGCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCCACC 5018

Db        826 --------------------------MetThrProAsnProGlyTyrGlnProSerMe 836

Qy       5019 AAAC----------------------AAGATGACCCAAAACTTCGATCCAATCTCTTC 5054

Db        836 tAsnThrSerAspMetMetGlyArgMetSerTyrGluProAsnLysAspProTyrGlySe 856

Qy       5055 TTTGGCACAAATGTCCCAACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGG 5114

Db        856 rMetArgLysAlaProGlySerAspProPheMetSerSerGlyGlnGlyProAsnGlyGl 876

Qy       5115 AACTGGTGGT---ATGACGATGATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCA 5171

Db        876 yMetGlyAspProTyrSerArgAlaAlaGlyProGlyLeuGlyAsnVal----------- 892

Qy       5172 TGGAATAATTTCGGGACTAGATGGATCAGGAATAGATACCATAAATCAAAATAACTGTCA 5231

Db        892 ------------------------------------------------------------ 892

Qy       5232 TTCAATGAATGTCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGT 5291

Db        893 ----------------------AlaMetGlyProArgGlnHisTyrPro----------- 901

Qy       5292 AGCAGGCGGTCCAAAT--------------------------GGACCG----------- 5313

Db        902 -TyrGlyGlyProTyrAspArgValArgThrGluProGlyIleGlyProGluGlyAsnMe 921

Qy       5314 -------------------CCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTA----- 5349

Db        921 tSerThrGlyAlaProGlnProAsnLeuMetProSerAsnProAspSerGlyMetTyrSe 941

Qy       5350 ---------------------------------------------AGAGAGAATTCCATAGG 5366

Db        941 rProSerArgTyrProProGlnGlnGlnGlnGlnGlnGlnArgHisAspSerTyrGl 961

Qy       5367 G---------TCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGG----------- 5406

Db        961 yAsnGlnPheSerThrGlnGlyThrProSerGlySerProPheProSerGlnGlnThrTh 981

Qy       5407 -------------------------------------------------GTTGTTCCACC 5417

Db        981 rMetTyrGlnGlnGlnGlnGlnAsnTyrLysArgProMetAspGlyThrTyrGlyProPr 1001

Qy       5418 TGGTGCCAGAATGATGGGTCGAATG---------CCAGTCAATTTTGGTTCGAATTTCAA 5468

Db       1001 oAlaLysArgHisGluGlyGluMetTyrSerValProTyrSerThrGlyGlnGlyGlnPr 1021

Qy       5469 TCCGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATAC------------AT 5516

Db       1021 oGlnGlnGlnGlnLeuProProAlaGlnProGlnProAlaSerGlnGlnGlnAlaAlaGl 1041

Qy       5517 GCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGGA----------------- 5559

Db       1041 nProSerProGlnGlnAspValTyrAsnGlnTyrGlyAsnAlaTyrProAlaThrAlaTh 1061

Qy       5560 -GCTAATAATGTGCGAATGCCA---------CCTAGTCTGGAATTTTTGCAGAGGTACGC 5609

Db       1061 rAlaAlaThrGluArgArgProAlaGlyGlyProGlnAsnGlnPheProPheGlnPheGl 1081

Qy       5610 TAACCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGA 5669

Db       1081 yArgAspArgValSerAla----------------------ProProGlyThrAsnAl 1093

Qy       5670 CGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCA---------------------- 5706

Db       1093 aGlnGlnAsnMetProProGlnMetMetGlyGlyProIleGlnAlaSerAlaGluValAl 1113

Qy       5707 ----------------------GGAGCCGGAGGTATGCTAATGAATTCTTCCGGAGAGCA 5744

Db       1113 aGlnGlnGlyThrMetTrpGlnGlyArgAsnAspMetThrTyrAsnTyrAlaAsnArgGl 1133

Qy       5745 ACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAA 5804

Db       1133 nSerThrGlySerAlaProGlnGlyPro---AlaTyrHisGlyValAsn---------Ar 1149

Qy       5805 TTGCAATCAAATGTCTATTGTTGACGAA------------GAGGGTGGATTACCCGGCCA 5852

Db       1149 gThrAspGluMetLeuHisThrAspGlnArgAlaAsnHisGluGlySerTrpProSerHi 1169

Qy       5853 TGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCAT 5912

Db       1169 s------------------------------------------------------------ 1169

Qy       5913 GCGGCCAAATGTAATGGGTGCGCGGATGCCACCC 5946

Db       1170 ----------------GlyThrArgGlnProPro 1175

RESULT 11
US-09-270-767-61668
; Sequence 61668, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61668
;   LENGTH: 43
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-270-767-61668

Alignment Scores:
Pred. No.:              1.11e-13        Length:         43
Score:                  236.00          Matches:        43
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            1.96%           Indels:         0
DB:                     4               Gaps:           0

US-10-664-859-1 (1-6909) x US-09-270-767-61668 (1-43)

Qy       2808 AAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATA 2867

Db          1 LysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerValGlyMetProTrpIle 20

Qy       2868 GGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAG 2927

Db         21 GlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLysIleThrGlnGlnGln 40

Qy       2928 CCACATACA 2936

Db         41 ProHisThr 43

RESULT 12
US-09-949-016-7421
; Sequence 7421, Application US/09949016
```

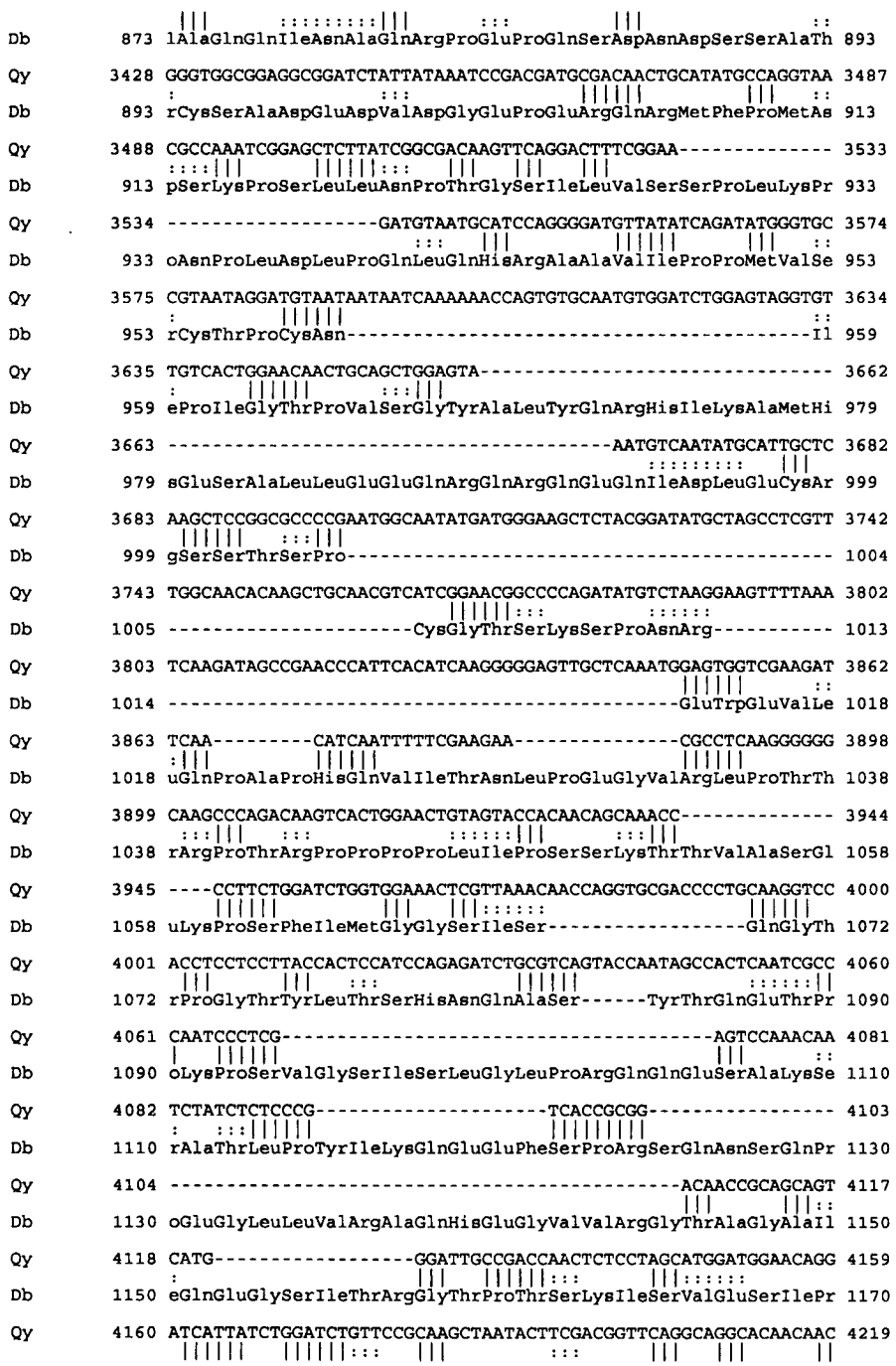
```
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7421
;   LENGTH: 1076
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7421

Alignment Scores:
Pred. No.:              5.2e-12       Length:          1076
Score:                  227.00        Matches:         247
Percent Similarity:     31.89%        Conservative:    158
Best Local Similarity:  19.45%        Mismatches:      419
Query Match:            1.89%         Indels:          446
DB:                     4             Gaps:            60

US-10-664-859-1 (1-6909) x US-09-949-016-7421 (1-1076)

Qy       2883 GGACTAACTCCT---------CCTAATCCTGTAGCCAAAATAACACAACAGCAGCCACAT 2933
Db          5 GlyLeuSerProGlnAspAspThrAsnSerGlyMetSerIleProArgValAsnProSer 24

Qy       2934 ACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAACAGCAAACGTAGT 2993
Db         25 -----------------ValAsnProSerIleSerProAlaHisGlyValAlaArgSer 38

Qy       2994 ACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGATCCTATGGGCAACGAAACT 3053
Db         39 SerThrLeuProProSerAsnSerAsnMet---------------------------- 48

Qy       3054 GAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAACTCACGA 3113
Db         49 ---------------------ValSerThrArgIleAsnArgGln---GlnSerSerAsp 60

Qy       3114 AATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAG---------------------- 3152
Db         61 LeuHisSerSerSerHisSerAsnSerSerAsnSerGlnGlySerPheGlyCysSerPro 80

Qy       3153 GCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTC---ACAAAAGGAAGC 3209
Db         81 GlySerGlnIleValAlaAsnValAlaLeuAsnGlnGlyGlnAlaSerSerGlnSerSer 100

Qy       3210 GATCCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGAC 3269
Db        101 AsnProSerLeuAsnLeuAsnAsnSerProMetGluGlyThrGlyIleSerLeuAla--- 119

Qy       3270 GAAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAA 3329
Db        120 GlnPheMetSerPro------ArgArgGlnValThrSerGlyLeuAlaThrArgProArg 137

Qy       3330 ATGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCT--------------- 3374
Db        138 MetProAsnAsnSerPheProProAsnIleSerThrLeuSerSerProValGlyMetThr 157

Qy       3375 --------------AATGTAAGCTCACAGATAACAAAAATTCCAGGAGATTTAATGATG 3419
Db        158 SerSerAlaCysAsnAsnAsnAsnArgSerTyrSerAsnIleProValThrSerLeuGln 177

Qy       3420 GGGATGTCGGGTGGCGGAGGCGGATCTATT------------ATAAATCCGACGATGCGA 3467

Db        178 GlyMetAsnGluGlyProAsnAsnSerValGlyPheSerAlaSerSerProValLeuArg 197

Qy       3468 CAACTGCAT----------------------------ATGCCAGGTAACGCCAAATCG 3497
Db        198 GlnMetSerSerGlnAsnSerProSerArgLeuAsnIleGlnProAlaLysAlaGluSer 217

Qy       3498 -------------------------------GAGCTCTTATCGGCGACAAGTTCAGGA 3524
Db        218 LysAspAsnLysGluIleAlaSerIleLeuAsnGluMetIleGlnSerAspAsnSerSer 237

Qy       3525 --------------CTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATG 3569
Db        238 SerAspGlyLysProLeuAspSerGlyLeuLeuHisAsnAsnAspArgLeuSerAsp--- 256

Qy       3570 GGTGCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTA 3629
Db        257 -----------GlyAspSerLysTyrSerGlnThrSerHisLysLeu---------Val 269

Qy       3630 GGTGTTGTCACTGGAACAACTGCA-----------GCTGGAGTAAATGTCAATATGCAT 3677
Db        270 GlnLeuLeuThr---ThrThrAlaGluGlnGlnLeuArgHisAlaAspIleAspThrSer 288

Qy       3678 TGC-----------TCAAGCTCCGGCGCCCCGAATGGCAATATGATGGGAAGCTCTACG 3725
Db        289 CysLysAspValLeuSerCysThrGlyThrSerAsnSerAlaSerAlaAsnSerSerGly 308

Qy       3726 GATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATCGGAACGGCCCCAGATATG 3785
Db        309 ----------------------------------------GlySerCysProSer--- 313

Qy       3786 TCTAAGGAAGTTTTAAATCAAGATAGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAA 3845
Db        314 --------------SerHisSerSerLeuThrGluArgHis----------------- 322

Qy       3846 ATGGAGTGGTCGAAGATTCAACATCAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCC 3905
Db        323 -----------LysIleLeuHisArgLeuLeuGlnGlu-----------GlySerPro 334

Qy       3906 AGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACCCCTTCTGGATCTGGTGGAAAC 3965
Db        335 SerAspIleThrThrLeuSerValGluProAspLysLysAspSerAlaSerThrSerVal 354

Qy       3966 TCGTTAAACAACCAGGTG---------------------------------------- 3983
Db        355 SerValThrGlyGlnValGlnGlyAsnSerSerIleLysLeuGluLeuAspAlaSerLys 374

Qy       3984 ----------------------------------------------------CGACCC 3989
Db        375 LysLysGluSerLysAspHisGlnLeuLeuArgTyrLeuLeuAspLysAspGluLysAsp 394

Qy       3990 CTGCAAGGTCCACCT--------------CCTCCTTACCACTCCATCCAGAGATCTGCG 4034
Db        395 LeuArgSerThrProAsnLeuSerLeuAspAspValLysValLysValGluLysLysGlu 414

Qy       4035 TCAGTACCAATAGCCACTCAATCGCCCAATCCC---------TCGAGTCCAAACAATCTA 4085
Db        415 GlnMetAspProCysAsnThrAsnProThrProMetThrLysProThrProGluGluIle 434

Qy       4086 TCTCTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGA--------------TTGCCG 4130
Db        435 LysLeuGluAlaGlnSerGlnPheThrAlaAspLeuAspGlnPheAspGlnLeuLeuPro 454

Qy       4131 ACC--------------AACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGATCT 4175
Db        455 ThrLeuGluLysAlaAlaGlnLeuProGlyLeuCysGluThrAspArgMetAspGlyAla 474

Qy       4176 GTTCCGCAAGCTAAT-----------------ACTTCGACGGTTCAGGCAGGCACAACA 4217
Db        475 ValThrSerValThrIleLysSerGluIleLeuProAlaSerLeuGlnSerAlaThrAla 494

Qy       4218 ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAA 4277
```

```
Db        495 ArgProThrSerArgLeuAsnArgLeuProGluLeuGluLeuGluAlaIleAspAsnGln 514

Qy       4278 AATCGTAGTAGAAATACCGGA----------------TCGTCAAGCGTTCTTACGCAT 4319
                              ||||||              :::::::::|||
Db        515 PheGlyGlnProGlyThrGlyAspGlnIleProTrpThrAsnAsnThrValThrAlaIle 534

Qy       4320 AACTTAAGCAGCAACCCA-----------AGTACCCCCTTATCTCATCTA------TCC 4361
              |||   |||   :::                |||:::   |||       |||
Db        535 AsnGlnSerLysSerGluAspGlnCysIleSerSerGlnLeuAspGluLeuLeuCysPro 554

Qy       4362 CCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTT 4421
              |||         |||                  |||   :::      |||   :::|||
Db        555 ProThrThrValGluGlyArgAsnAspGluLysAlaLeuLeuGluGlnLeuValSerPhe 574

Qy       4422 TTT------AAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAA 4475
                    ||||||::::::   :::      ::::::      |||   |||
Db        575 LeuSerGlyLysAspGluThrGluLeuAlaGluLeuAspArgAlaLeuGlyIleAspLys 594

Qy       4476 CTC------GGAAATTTGATAGAAAAAATCAGGAATAGA---------------------- 4508
              |||      |||     ::::::   :::      |||
Db        595 LeuValGlnGlyGlyGlyLeuAspValLeuSerGluArgProProGlnGlnAlaThr 614

Qy       4509 ---------------------AAAAATAAATTATTTTCC-----------------GGA 4529
                                  :::      |||:::|||
Db        615 ProProLeuIleMetGluGluArgProAsnLeuTyrSerGlnProTyrSerSerProSer 634

Qy       4530 CCGCCCATCCATTTCTTGAATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTA 4589
              |||       :::      :::|||      |||::::::   :::          |||
Db        635 ProThrAlaAsnLeuProSerProPheGlnGlyMetValArgGlnLysProSerLeuGly 654

Qy       4590 AAATTAAACACGAAAATTCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATT 4649
                 :::      ::::::     |||
Db        655 ThrMetProValGlnVal--ThrPro---------------------------------- 662

Qy       4650 AAAAATATGTTTATTATAATATTTCTACAGGTGATAACATGAAAAGTAGG---------- 4699
                                  |||   |||      |||::   |||
Db        663 --------ProArgGlyAlaPheSerProGlyMetGlyMetGlnProArgGlnThrLeuA 680

Qy       4700 --CGACCA------AGCCCA---------------------------------CAGGGTC 4718
                |||||      :::|||                                 |||||||
Db        680 snArgProProAlaAlaProAsnGlnLeuArgLeuGlnLeuGlnGlnArgLeuGlnGlyG 700

Qy       4719 AGCGGTCACCAGTAAATAGTCTAATAGAGGCAAATAAAGATGTA---------CGATTTG 4769
              ||:::              |||||    |||:::                 :::||||
Db        700 lnGln------------GlnLeuIleHisGlnAsnArgGlnAlaIleLeuAsnGlnPheA 716

Qy       4770 CTGCATCCAGTCCT---GGTTTTAACCCGCATCCACATATGCAAAGC------AATTCAA 4820
              |||||::::::|||   |||   |||            ||||||
Db        716 laAlaThrAlaProValGlyIleAsnMetArgSerGlyMetGlnGlnGlnIleThrProG 736

Qy       4821 ATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAATATACAGATGGAGGTAAATATTT 4880
                      |||||||||      :::              ::::::
Db        736 lnProProLeuAsnAlaGlnMetLeuAlaGlnArgGlnArgGluLeu------------- 751

Qy       4881 AAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGC 4940
                                                 :::|||   :::|||::::
Db        752 ----------------------------------------TyrSerGlnGlnHisArgG 758

Qy       4941 GCAAGGTGGATCCGTACAATTTAGTCGGCGCTCCGATAATATTCCGCTAAATCCCAATAG 5000
              :: |||   :::
Db        758 ln-ArgGlnLeu------------------------------------------------ 761

Qy       5001 TGGCAATCGGCCGCCACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGC 5060
                           :::|||              |||                |||||
Db        762 --------------IleGlnGlnGlnArgAlaMetLeuMetArgGlnGlnSerPheGly 776

Qy       5061 ACAAATGTCCCAACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGG 5120
                 |||:::|||   ::::::   |||   |||||| ||||||
Db        777 AsnAsnLeuProProSerSerGlyLeuProValGln-MetGly----------------- 790

Qy       5121 TGGTATGACGATGATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCATGGAATAAT 5180

Db        790 ------------------------------------------------------------ 790

Qy       5181 TTCGGGACTAGATGGATCAGGAATAGATACCATAAATCAAAATAACTGTCATTCAATGAA 5240

Db        790 ------------------------------------------------------------ 790

Qy       5241 TGTCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGG 5300
                                                 ||||||::::::
Db        791 --------------------------------AsnProArgLeu-------------- 794

Qy       5301 TCCAAATGGACCGCCT-------------------------GGCTTTAATCCTAATTC 5333
              |||   |||   |||                        |||   ||||||   ::
Db        795 -ProGlnGlyAlaProGlnGlnPheProTyrProProAsnTyrGlyThrAsnProGlyTh 814

Qy       5334 CCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAACTCTTC 5393
              :|||                  |||      :::          |||:::   ::
Db        814 rProProAlaSerThrSerProPheSerGlnLeuAlaAlaAsnProGluAlaSerLeuAl 834

Qy       5394 AAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTCAATTT 5453
              :|||         :::|||      |||      |||   |||   :::          ||
Db        834 aAsnArgAsnSerMetValSerArgGly------MetThrGlyAsnIleGlyGlyGlnPh 852

Qy       5454 TGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATA 5513
              ||||:::      ||||||  :::|||                |||      |||||
Db        852 eGlyThrGlyIleAsnProGlnMetGln--------------GlnAsnValPheGlnTy 867

Qy       5514 C---------------ATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGG 5558
              |               :::|||            |||          :::   :::   ||
Db        867 rProGlyAlaGlyMetValProGlnGlyGluAlaAsnPheAlaProSerLeuSerProGl 887

Qy       5559 AGCTAATAATGTGCGAATG---------CCACCTAGTCTGGAATTTTTGCAGAGG----- 5604
              |::::::   |||   |||          ||||||          ||||||:::
Db        887 ySerSerMetValProMetProIleProProProGlnSerSerLeuLeuGlnGlnThrPr 907

Qy       5605 ------------TACGCTAACCCT--------------CAAATGGGTGCTGTAGGCAA 5636
                         |||   :::|||              |||   ||||||:::|||||
Db        907 oProAlaSerGlyTyrGlnSerProAspMetLysAlaTrpGlnGlnGlyAlaIleGlyAs 927

Qy       5637 TGGGTCGCCAATATGCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCCAGGATTG-- 5694
              |   :::                  :::      ||||||      ||||||:::
Db        927 nAsnAsnValPheSerGlnAlaValGlnAsnGlnProThrProAlaGlnProGlyValTy 947

Qy       5695 ------------------------ATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAA 5729
                                      |||||||||                  |||||
Db        947 rAsnAsnMetSerIleThrValSerMetAlaGlyGlyAsnThrAsnValGlnAsnMetAs 967

Qy       5730 TTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGC-------- 5781
              |         |||   |||                |||||   :::
Db        967 nProMetMetAlaGlnMetGlnMetSerSerLeuGlnMetProGlyMetAsnThrValCy 987

Qy       5782 ---------------------------------AATGGTATTAACTTCTTTCAGAATTG 5807
                                               |||:::               ||
Db        987 sProGluGlnIleAsnAspProAlaLeuArgHisThrGlyLeu------------TyrCy 1003

Qy       5808 CAATCAAATGTCTATTGTTGAC---------GAAGAGGGTGGATTACCCGGCCATGACGG 5858
              |||||||:::|||      |||          |||                |||||
Db       1003 sAsnGlnLeuSerSerThrAspLeuLeuLysThrGluAla--------------AspGl 1018

Qy       5859 ATCAATGAATATTGGTCAACCATCTATG--------------------------ATAAG 5891
              |:::      :::   |||   :::                           :::
Db       1018 yThrGlnGlnValGlnGlnValGlnValPheAlaAspValGlnCysThrValAsnLeuVa 1038

Qy       5892 GGGCATGCGTCCACATGCCATGCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGTTAA 5951
                |||      |||:::      :::|||      :::|||   :::   |||
Db       1038 lGlyGlyAspProTyrLeuAsnGlnProGlyProLeuGlyThrGlnLysProThrSerGl 1058

Qy       5952 CAGGCAAATTCAGTTTGCACAG 5973
                  |||          ||||||
Db       1058 yProGlnThrProGlnAlaGln 1065
```

RESULT 13
US-09-125-635-4
; Sequence 4, Application US/09125635
; Patent No. 6562589

```
; GENERAL INFORMATION:
;  APPLICANT:  THE UNITED STATES OF AMERICA represented by THE SE
;  TITLE OF INVENTION:  AIB1, A novel steriod receptor co-activator
;  FILE REFERENCE:  49944
;  CURRENT APPLICATION NUMBER: US/09/125,635
;  CURRENT FILING DATE:  1998-08-21
;  PRIOR APPLICATION NUMBER:  60/049,728
;  PRIOR FILING DATE:  1997-06-17
;  NUMBER OF SEQ ID NOS:  12
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 1420
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-125-635-4

Alignment Scores:
Pred. No.:              1.81e-11      Length:         1420
Score:                  222.00        Matches:        267
Percent Similarity:     30.04%        Conservative:   150
Best Local Similarity:  19.24%        Mismatches:     496
Query Match:            1.84%         Indels:         475
DB:                     4             Gaps:           65

US-10-664-859-1 (1-6909) x US-09-125-635-4 (1-1420)

Qy       3120 GTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTGGAAGCAGCTGGC 3179
Db         95 ValGlnLysAlaAspValSerSerThrGlyGlnGlyValIleAspLysAspSerLeuGly 114

Qy       3180 GTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACTGAAAACAACATT 3239
Db        115 ProLeuLeuLeuGlnAlaLeuAspGlyPheLeuPheValValAsnArgAspGlyAsnIle 134

Qy       3240 GTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAACAGCGGCAACAT 3299
Db        135 ValPheVal----------------SerGluAsnValThr---GlnTyrLeuGlnTyr 147

Qy       3300 CGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTTCCTGAAAATGAG 3359
Db        148 LysGlnGluAspLeuValAsnThrSerValTyrAsn---IleLeuHisGluGluAspArg 166

Qy       3360 AATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGAGATTTAATGATG 3419
Db        167 LysAspPhe----------------LeuLysAsnLeuProLysSerThrValAsn 179

Qy       3420 GGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGACAACTGCATATG 3479
Db        180 GlyValSerTrp----------------ThrAsnGluThrGlnArgGlnLysSerHis 193

Qy       3480 CCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTTTCGGAAGATGTA 3539
Db        194 ThrPheAsnCysArg------MetLeuMetLysThr---------------------- 203

Qy       3540 ATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGTAATAATAATCAA 3599
Db        204 ------ProHisAspIleLeuGluAspIleAsnAlaSerProGluMetArgGlnArgTyr 221

Qy       3600 AAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACAACTGCAGCTGGA 3659
Db        222 GluThr---------------------------------------------------- 223

Qy       3660 GTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAATATGATGGGAAGC 3719
Db        224 ------------MetGlnCysPheAlaLeuSerGlnProArgAla---MetMetGluGlu 238

Qy       3720 TCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTC-------------- 3764
Db        239 GlyGluAspLeu----------------GlnSerCysMetIleCysValAlaArgArg 252

Qy       3765 ----------------------------------------ATCGGAACGGCCCCAGAT 3782
Db        253 IleThrThrGlyGluArgThrProProSerAsnProGluSerPheIleThrArgHisAsp 272

Qy       3783 ATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACC------CATTCACATCAAGGGGGA 3836
Db        273 LeuSerGlyLysValValAsnIleAspThrAsnSerLeuArgSerSerMetArgProGly 292

Qy       3837 GTTGCTCAAATG-------------------------------------------GAG 3851
Db        293 PheGluAspIleIleArgArgCysIleGlnArgPhePheSerLeuAsnAspGlyGlnSer 312

Qy       3852 TGGTCGAAGATTCAACATCAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAA 3911
Db        313 TrpSerGlnLysArgHis---TyrGlnGluAlaTyrLeuAsnGlyHisAlaGluThrPro 331

Qy       3912 GTC-----------------ACTGGAACTGTAGTACCACAACAGCAAACCCCTTCTGGA 3953
Db        332 ValTyrArgPheSerLeuAlaAspGlyThrIleVal---ThrAlaGlnThrLysSerLys 350

Qy       3954 TCTGGTGGAAACTCGTTAAACAACCAG------------------------------- 3980
Db        351 LeuPheArgAsnProValThrAsnAspArgHisGlyPheValSerThrHisPheLeuGln 370

Qy       3981 --------------GTGCGACCCCTG--------------CAAGGTCCACCTCCTCCT 4010
Db        371 ArgGluGlnAsnGlyTyrArgProAsnProAsnProValGlyGlnGlyIleArgProPro 390

Qy       4011 TACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACTCAATCGCCCAAT------ 4064
Db        391 MetAlaGlyCys------AsnSerSerValGlyGlyMetSerMetSerProAsnGlnGly 408

Qy       4065 ---------CCCTCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCAGCA 4115
Db        409 LeuGlnMetProSerSer---ArgAlaTyrGlyLeuAlaAspProSerThrThrGlyGln 427

Qy       4116 GTCATGGGATTGCCG-----------------------------ACCAACTCTCCTAGC 4145
Db        428 MetSerGlyAlaArgTyrGlyGlySerSerAsnIleAlaSerLeuThrProGlyProGly 447

Qy       4146 ATGGATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAG 4205
Db        448 MetGlnSerProSerSerTyrGlnAsnAsnAsnTyrGlyLeuAsnMetSerSerProPro 467

Qy       4206 GCAGGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCG 4265
Db        468 HisGlySerProGlyLeu-----------------------------------------Ala 474

Qy       4266 CCGTCAAATCAAAAT---------CGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACG 4316
Db        475 ProAsnGlnGlnAsnIleMetIleSerProArgAsnArgGlySerProLysIleAlaSer 494

Qy       4317 CATAACTTAAGC------AGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAGGAA 4370
Db        495 HisGlnPheSerProValAlaGlyValHisSerProMetAlaSerSerGlyAsnThrGly 514

Qy       4371 TTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAAT------------ 4418
Db        515 AsnHisSerPheSerSerSerSerLeuSerAlaLeuGlnAlaIleSerGluGlyValGly 534

Qy       4419 -----------------------------------TTTTTAAAGACAAATCAAATATG 4442
Db        535 ThrSerLeuLeuSerThrLeuSerSerProGlyProLysLeuAspAsnSerProAsnMet 554

Qy       4443 AATTGCGTTAATAATAAGTTATATATTACATAACTCGGAAATTTGATAGAAAAAATCAGG 4502
Db        555 AsnIleThrGlnPro---------------------------------SerLysValSer 563

Qy       4503 AATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCATTTCTTG---------AATCCA 4553
Db        564 AsnGlnAspSerLys------------SerProLeuGlyPheTyrCysAspGlnAsnPro 579

Qy       4554 ATTTCTGGAGTGATTGTT------------AGAGATAATCTACTATTAAAA--------- 4592
Db        580 ValGluSerSerMetCysGlnSerAsnSerArgAspHisLeuSerAspLysGluSerLys 599
```

```
Db       1220 nMetValAlaGlnArgSerArgGluLeuLeuSerHisHisPheArgGlnGlnArgValAl 1240

Qy       6304 ----AAGTATGCCCAGCAATATCATAGTTTTCAGCAGCAGTTATATGCTACCAACACCAG 6359
                  ||||||           |||||||||                          ::
Db       1240 aMetMetMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGl 1260

Qy       6360 AAGTCAACAACAACAGCATATGCACCAGCAGCACCAGAGCAACATGATAACAATGCCGCC 6419
              :  ||||||||||||          ||||||   |||:::         :::   |||||
Db       1260 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnThrGlnAlaPheSerProProPr 1280

Qy       6420 GAATTTATCACCAAATCCAACG 6441
              ||||::::::   :::|||:::
Db       1280 oAsnValThrAlaSerProSer 1287

RESULT 14
US-09-949-016-9675
; Sequence 9675, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9675
;   LENGTH: 2451
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-9675

Alignment Scores:
Pred. No.:              4.8e-11       Length:          2451
Score:                  219.00        Matches:         297
Percent Similarity:     31.80%        Conservative:    223
Best Local Similarity:  18.17%        Mismatches:      557
Query Match:            1.82%         Indels:          560
DB:                     4             Gaps:            75

US-10-664-859-1 (1-6909) x US-09-949-016-9675 (1-2451)

Qy       2249 ATATATAGAAGTCGCCTTCACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCC 2308
              :::|||:::   |||              :::|||         |||:::
Db        441 ValTyrLysAspArg-----------------GlnPheMetAsnValTrpThrAspHis 454

Qy       2309 CGCAGTGATTTATATAGA--------------------------------TGATTTACG 2335
                    ::::::::::::
Db        455 GluLysGluIlePheLysAspLysPheIleGlnHisProLysAsnPheGlyLeuIleAla 474

Qy       2336 AGTTATTTA-----------------------ATTTTTTATGGTGTATTTTAATAAATA 2371
              |||||||||                      :::   |||
Db        475 SerTyrLeuGluArgLysSerValProAspCysValLeuTyr------------------ 488

Qy       2372 TCTTATTTATTCATTTTACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGA 2431
                |||                 :::   |||   |||         |||:::   :::|||
Db        489 ---TyrTyrLeuThrLysLysAsnGluAsnTyrLysAlaLeuValArgArgAsnTyrGly 507

Qy       2432 AAAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCA 2491
              |||:::|||                       ||||||:::                |||
Db        508 LysArgArg--------------------GlyArgAsnGlnGlnIleAlaArgProSer 520

Qy       2492 GTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGT 2551
                |||   :::      :::         |||                 |||
Db        521 GlnGluGluLysValGluGlu---------LysGluGluAspLysAlaGluLysThrGlu 537

Qy       2552 AAGCGG-------------------------------------------AGTAGGACC 2566
              |||:::                                          |||:::
Db        538 LysLysGluGluGluLysLysAspGluGluGluLysAspGluLysGluAspSerLysGlu 557

Qy       2567 AATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGG 2626
              ||||||:::   :::           ||||||      :::   :::   :::
Db        558 AsnThrLysGluLysAsp---LysIleAspGlyThrAlaGluGluThrGluGluArg--- 575

Qy       2627 AATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCACATATTCGT 2686
                                                      ||||||   :::         :::
Db        576 ----------------------------GluGlnAlaThrProArgGlyArg---Lys 584

Qy       2687 GTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTTCAAAC 2746
                    ||||||   |||:::::::::|||||||:::      ||||||:::::|||
Db        585 ThrAlaAsnSerGlnGlyArgArgLysGlyArgIleThr------ArgSerMetThrAsn 602

Qy       2747 TATTATTGCGTATCACTGCACTCAGCCTGCTAC-----------------AAAAAGCTT 2788
                        :::      :::|||                              |||
Db        603 GluAlaAlaAlaAlaSerAlaAlaAlaAlaAlaAlaThrGluGluProProProProLeu 622

Qy       2789 CCTGGAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTC 2848
              |||                  :::|||          |||      :::::::::
Db        623 ProProPro---------ProGluProIleSerThrGluProValGluThrSerArgTrp 639

Qy       2849 CGTCGGTATGCCATGGAT--------------------AGGCATGGGGCAGGTTGG--- 2884
                        |||:::                     ||||||      |||
Db        640 ThrGluGluGluMetGluValAlaLysLysGlyLeuValGluHisGlyArgAsnTrpAla 659

Qy       2885 ----------------------ACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACA 2923
                                  |||   |||      |||
Db        660 AlaIleAlaLysMetValGlyThrLysSerGluAlaGlnCys---------LysAsnPhe 676

Qy       2924 GCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAA----- 2978
                    ||||||   |||      ::::::      :::||| ||||||:::
Db        677 TyrPheAsnTyrLysArgArgHisAsnLeuAspAsnLeuLeuGln-GlnHisLysGlnLy 696

Qy       2979 ----------------------------------------AACAGCAAACGTAGTAC 2995
                                                      |||      |||||
Db        696 sThrSerArgLysProArgGluGluArgAspValSerGlnCysGluSerValAlaSerTh 716

Qy       2996 TGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCT-----------GATCCTATGGG 3043
              |||||||||      :::      ::::::   |||           :::|||
Db        716 rValSerAlaGlnGluAspGluAspIleGluAlaSerAsnGluGluGluAsnProGluAs 736

Qy       3044 CAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACA 3103
                :::|||   |||   :::                |||         :::   |||
Db        736 pSerGluValGluAlaValLysProSerGluAspSerProGluAsnAlaThrSerArgGl 756

Qy       3104 AAACTCACGAAATCATGTA---------------------------------------- 3122
              |||:::         |||
Db        756 yAsnThrGluProAlaValGluLeuGluProThrThrGluThrAlaProSerThrSerPr 776

Qy       3123 ------------------------------------GACAGTATCAGTACATCC---AG 3142
                                                 :::|||:::   |||     ::
Db        776 oSerLeuAlaValProSerThrLysProAlaGluAspGluSerValGluThrGlnValAs 796

Qy       3143 CGAGTCACAGGCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAG------GT 3196
              ::::|||   :::   :::      |||         ||||||      :::
Db        796 nAspSerIleSerAlaGluThrAlaGluGlnMetAspValAspGlnGlnGluHisSerAl 816

Qy       3197 CACAAAAGGAAGC------GATCCTGGCCTGACAACTGAAAACAACATTGTATCACTGCA 3250
                ::::||||||      ||||||         |||:::   :::   |||   :::::
Db        816 aGluGluGlySerValCysAspProProProAlaThrLysAlaAspSerValAspValGl 836

Qy       3251 AGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACA 3310
              :  |||:::||||||:::   :::   :::       :::       :::   :::|||::
Db        836 u---ValArgValProGluAsnHisAlaSerLysValGluGlyAspAsnThrLysGluAr 855

Qy       3311 GTTGGCAAAAATAAAAAAAAATGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGG 3370
              :         :::   :::   ::::::   :::   |||   :::|||:::   |||
Db        855 g------AspLeuAspArgAlaSerGluLysValGluProArgAspGluAspLeuValVa 873

Qy       3371 AGCT---AATGTAAGCTCACAGATAACAAAAATTCCAGGAGATTTAATGATGGGGATGTC 3427
```


Db 873 lAlaGlnGlnIleAsnAlaGlnArgProGluProGlnSerAspAsnAspSerSerAlaTh 893
Qy 3428 GGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGACAACTGCATATGCCAGGTAA 3487
Db 893 rCysSerAlaAspGluAspValAspGlyGluProGluArgGlnArgMetPheProMetAs 913
Qy 3488 CGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTTTCGGAA---------- 3533
Db 913 pSerLysProSerLeuLeuAsnProThrGlySerIleLeuValSerSerProLeuLysPr 933
Qy 3534 ----------GATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGC 3574
Db 933 oAsnProLeuAspLeuProGlnLeuGlnHisArgAlaAlaValIleProProMetValSe 953
Qy 3575 CGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGT 3634
Db 953 rCysThrProCysAsn----------Il 959
Qy 3635 TGTCACTGGAACAACTGCAGCTGGAGTA---------- 3662
Db 959 eProIleGlyThrProValSerGlyTyrAlaLeuTyrGlnArgHisIleLysAlaMetHi 979
Qy 3663 ----------AATGTCAATATGCATTGCTC 3682
Db 979 sGluSerAlaLeuLeuGluGluGlnArgGlnArgGlnGluGlnIleAspLeuGluCysAr 999
Qy 3683 AAGCTCCGGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTT 3742
Db 999 gSerSerThrSerPro---------- 1004
Qy 3743 TGGCAACACAAGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAA 3802
Db 1005 ----------CysGlyThrSerLysSerProAsnArg---------- 1013
Qy 3803 TCAAGATAGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGAT 3862
Db 1014 ----------GluTrpGluValLe 1018
Qy 3863 TCAA---------CATCAATTTTTCGAAGAA----------CGCCTCAAGGGGGG 3898
Db 1018 uGlnProAlaProHisGlnValIleThrAsnLeuProGluGlyValArgLeuProThrTh 1038
Qy 3899 CAAGCCCAGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACC---------- 3944
Db 1038 rArgProThrArgProProProProLeuIleProSerSerLysThrThrValAlaSerGl 1058
Qy 3945 ----CCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCC 4000
Db 1058 uLysProSerPheIleMetGlyGlySerIleSer----------GlnGlyTh 1072
Qy 4001 ACCTCCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACTCAATCGCC 4060
Db 1072 rProGlyThrTyrLeuThrSerHisAsnGlnAlaSer------TyrThrGlnGluThrPr 1090
Qy 4061 CAATCCCTCG----------AGTCCAAACAA 4081
Db 1090 oLysProSerValGlySerIleSerLeuGlyLeuProArgGlnGlnGluSerAlaLysSe 1110
Qy 4082 TCTATCTCTCCCG----------TCACCGCGG---------- 4103
Db 1110 rAlaThrLeuProTyrIleLysGlnGluGluPheSerProArgSerGlnAsnSerGlnPr 1130
Qy 4104 ----------ACAACCGCAGCAGT 4117
Db 1130 oGluGlyLeuLeuValArgAlaGlnHisGluGlyValValArgGlyThrAlaGlyAlaIl 1150
Qy 4118 CATG----------GGATTGCCGACCAACTCTCCTAGCATGGATGGAACAGG 4159
Db 1150 eGlnGluGlySerIleThrArgGlyThrProThrSerLysIleSerValGluSerIlePr 1170
Qy 4160 ATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGCACAACAAC 4219
Db 1170 oSerLeuArgGlySerIleThrGlnGlyThrProAlaLeuProGlnThrGlyIleProTh 1190
Qy 4220 AGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAAAA 4279
Db 1190 r---------- 1190
Qy 4280 TCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGCAGCAACCCAAG 4339
Db 1191 -GluAlaLeuValLysGlySerIleSerArgMet---------- 1201
Qy 4340 TACCCCCTTATCTCATCTATCCCCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGG 4399
Db 1202 ----ProIleGluAspSerSerProGluLysGlyArgGluGluAlaAlaSerLysGlyHi 1220
Qy 4400 TATGTTATATTTGTTTAATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAA 4459
Db 1220 sValIleTyr----------GluGlyLysSer---------- 1227
Qy 4460 GTTATATATTACATAACTCGGAAATTTGATA------GAAAAAATCAGGAATAGAAAAAA 4513
Db 1228 ----------GlyHisIleLeuSerTyrAspAsnIleLysAsnAlaArgGl 1241
Qy 4514 TAAATTATTTTC----CGGACCGCCCATCCATTTCTTGAATCCAATTTCTGGAGTGATTG 4569
Db 1241 uGlyThrArgSerProArgThrAlaHis---------- 1250
Qy 4570 TTAGAGATAATCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAATCA 4629
Db 1251 ----------GluIleSerLeuLysArgSerTyrGluSerValGluGly-- 1263
Qy 4630 CTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTACAGGTGATAACAT 4689
Db 1264 ----------AsnIleLysGlnGlyMetSerMe 1271
Qy 4690 GAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAGGC 4749
Db 1271 tArg----------GluSerProValSerAlaProLeuGluGl 1282
Qy 4750 AAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTTTTAACCCGCATCCACATATGCA 4809
Db 1282 y----------LeuIleCysArgAlaLeuProArgGlySerProHisSerAspLeuLy 1298
Qy 4810 AAGCAAT---------TCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAA--- 4857
Db 1298 sGluArgThrValLeuSerGlySerIleMetGlnGlyThrProArgAlaThrThrGluSe 1318
Qy 4858 -TATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCT 4916
Db 1318 rPheGluAspGlyLeuLysTyrProLysGlnIle---------- 1329
Qy 4917 TCTTTTTCAGCGTCAAGCA---------TCAGCGCAAGGTGGATCCGTACAATTTAGTCG 4967
Db 1330 -------LysArgGluSerProProIleArgAlaPheGluGlyAlaIleThrLysGlyLy 1347
Qy 4968 GCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGG----------CCGCC 5015
Db 1347 sProTyrAspGlyIleThrThrIleLysGluMetGlyArgSerIleHisGluIleProAr 1367
Qy 5016 ACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACA 5075
Db 1367 gGlnAspIleLeuThrGlnGluSerArgLysThrProGluValValGlnSerThrArgPr 1387
Qy 5076 ACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCA----------GCCGGA-- 5115
Db 1387 oIleIleGluGlySerIleSerGlnGlyThrProIleLysPheAspAsnAsnSerGlyGl 1407
Qy 5116 ----------ACTGGTGGTATGACGATGATGGGGGG 5141
Db 1407 nSerAlaIleLysHisAsnValLysSerLeuIleThrGlyProSerLysLeuSerArgGl 1427
Qy 5142 TCCGGGACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGATGGATCAGG 5201
Db 1427 yMetProProLeuGluIle---ValProGluAsnIleLysValValGluArgGlyLysTy 1446

```
Qy       5202 AATAGATACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAACTCAATGGG 5261
Db       1446 rGluAspValLysAlaGlyGluThrValArgSerArgHisThrSerValValSerSerGl 1466

Qy       5262 TCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACCGCCTGGCTT 5321
Db       1466 yProSerValLeuArgSerThrLeuHisGluAlaProLysAlaGlnLeuSerProGlyIl 1486

Qy       5322 TAATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGGCTGTGGCTC 5381
Db       1486 eTyrAspAspThr-----------------------------------------------Se 1491

Qy       5382 AGCAAACTCTTCA------AACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGG 5435
Db       1491 rAlaArgArgThrProValSerTyrGlnAsnThrMetSerArgGlySerProMetMetAs 1511

Qy       5436 TCGA---ATGCCAGTCAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAG 5492
Db       1511 nArgThrSerAspValThrIleSerSerAsnLysSerThrAsnHisGluArgLysSerTh 1531

Qy       5493 T---ACCCCAAACACCATACAATACATGCCAGTAAGGGCACAGAACGCCAAC-------- 5541
Db       1531 rLeuThrProThrGlnArgGluSerIleProAlaLysSerProValProGlyValAspPr 1551

Qy       5542 --------------------------AACAATAACAACAATGGAGCTAATAATGTG-- 5571
Db       1551 oValValSerHisSerProPheAspProHisHisArgGlySerThrAlaGlyGluValTy 1571

Qy       5572 -------------------CGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAA 5612
Db       1571 rArgSerHisLeuProThrHisLeuAspProAlaMetProPhe---HisArgAlaLeuAs 1590

Qy       5613 CCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGG 5672
Db       1590 pProAlaAlaAlaAlaTyrLeuPheGlnArgGlnLeuSerPro----------------- 1604

Qy       5673 TACTCCTGGAATGCCAGGA---------TTGATGGCGGGACCAGGAGCCGGAGGTATGCT 5723
Db       1605 -ThrProGlyTyrProSerGlnTyrGlnLeuTyrAlaMetGluAsnThrArgGlnThrIl 1624

Qy       5724 AATGAATTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAA 5783
Db       1624 eLeuAsnAspTyrIleThrSerGlnGlnMetGlnValAsnLeuArgProAspValAlaAr 1644

Qy       5784 TGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGTGGATT 5843
Db       1644 gGly----------------------------LeuSerProArgGluGlnProLeuGlyLe 1655

Qy       5844 ACCCGGCCATGACGGATCAATGAATATTGGTCAACCATCTATG---------------AT 5888
Db       1655 uPro----------------------------TyrProAlaThrArgGlyIleIleAspLe 1666

Qy       5889 AAGGGGCATGCGTCCACATGCCATGCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGT 5948
Db       1666 uThrAsnMetProProThrIleLeuValProHisProGlyGlyThrSerThrProProMe 1686

Qy       5949 TAACAGG--------------------CAAATTCAGTTTGCACAGTCATCGGATGGTAT 5987
Db       1686 tAspArgIleThrTyrIleProGlyThrGlnIleThrPhe------------------- 1699

Qy       5988 TGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTGCAACAGCGCTGGACC 6047
Db       1700 ---------------ProProArgProTyrAsnSerAlaSerMetSerProGlyHisPr 1714

Qy       6048 A---CACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACAACACATAAAGAA 6104
Db       1714 oThrHisLeuAlaAlaAlaAlaSerAlaGluArgGluArgGluArgGluArgGluLysGl 1734

Qy       6105 CATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACAAGGGCAGATCCA 6164
Db       1734 u-----------------ArgGluArgGluArgIleAlaAlaAlaSerSerAspLeuTy 1748

Qy       6165 ACTGCAT----------------------------GGGCAAGGACATGCGCAGGGTCA 6194
Db       1748 rLeuArgProGlySerGluGlnProGlyArgProGlySerHisGlyTyrValArgSerPr 1768

Qy       6195 GTCTTTAATTGGACCTACTAATAATAATTTAATGTCAACTGCCGGAAGTGTC---AGTGC 6251
Db       1768 oSerProSerValArgThrGlnGluThrMetLeuGlnGlnArgProSerValPheGlnGl 1788

Qy       6252 TACTAACGGTGTCTCTGGCATCAATTTCGTAGGTCCCTCTTCT 6294
Db       1788 yThrAsnGlyThrSerValIleThrProLeuAspProThrAla 1802
```

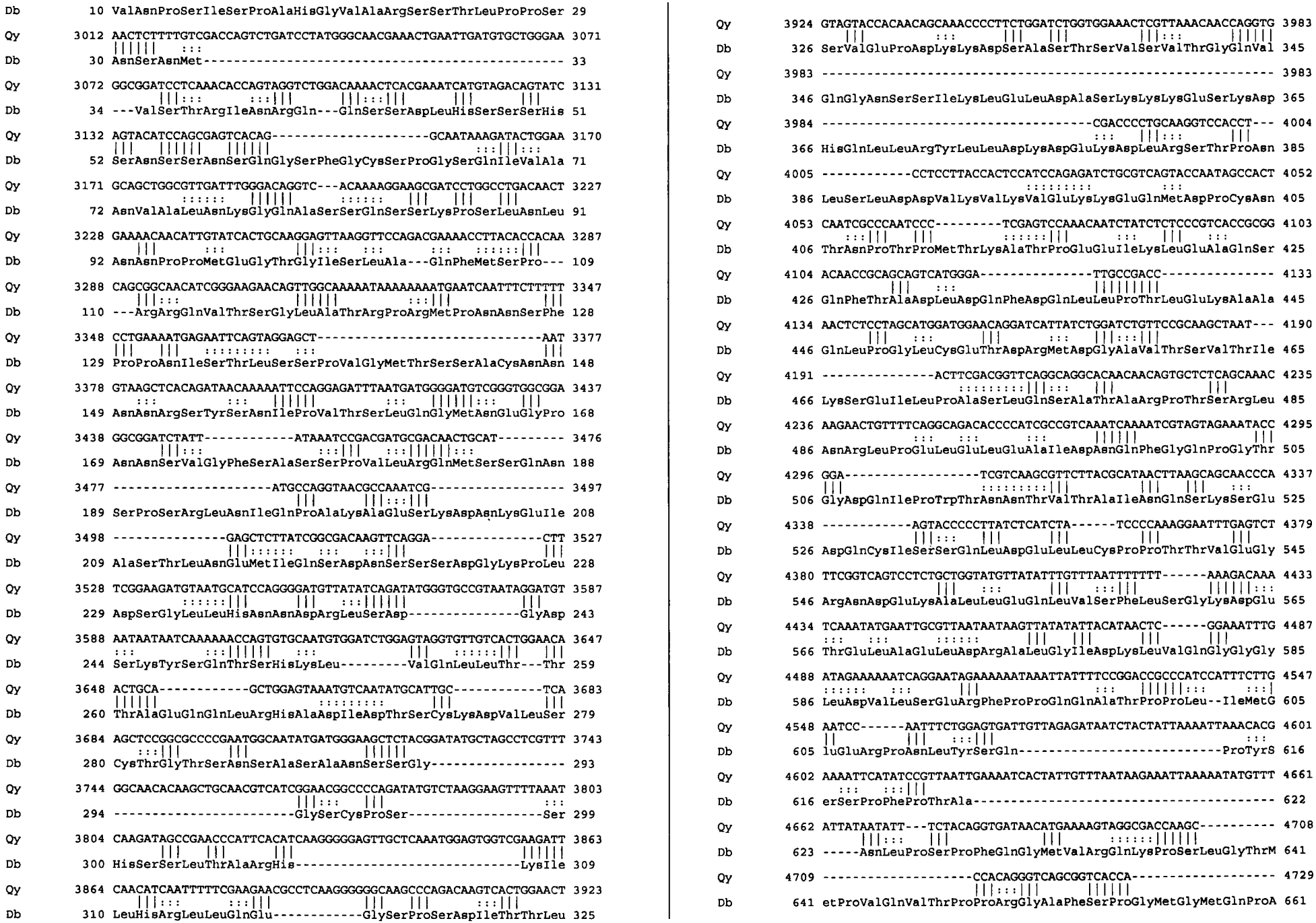
RESULT 15
US-08-701-154A-5
; Sequence 5, Application US/08701154A
; Patent No. 6380373
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Bert W.
; APPLICANT: Tsai, Ming-Jer
; APPLICANT: Tsai, Sophia Y.
; APPLICANT: Onate, Sergio A.
; TITLE OF INVENTION: STEROID RECEPTOR COACTIVATOR
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/701,154A
; FILING DATE: August 21, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/003,784
; FILING DATE: September 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 220/243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1061 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-701-154A-5

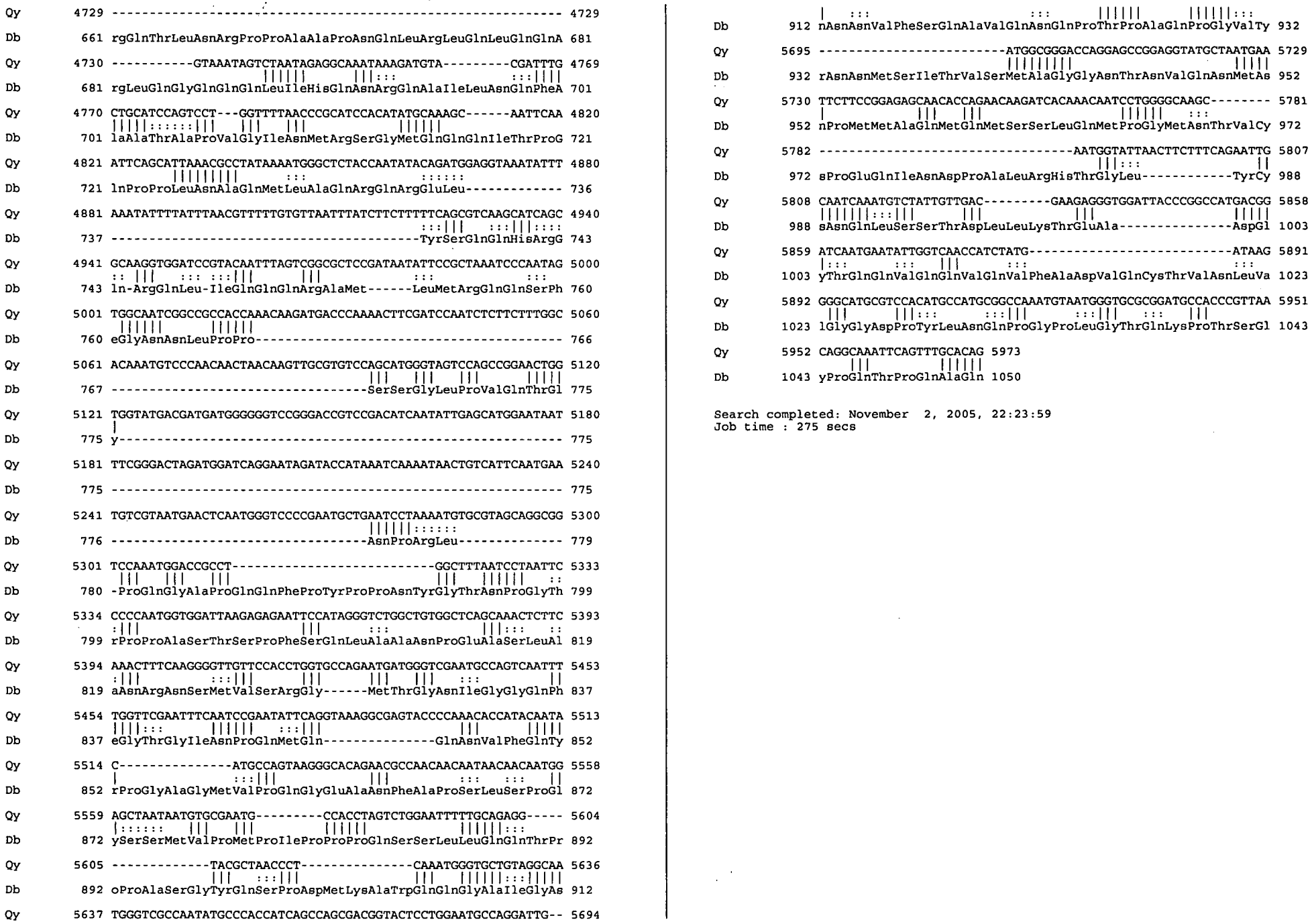
Alignment Scores:

Pred. No.:	5.05e-11	Length:	1061
Score:	216.50	Matches:	239
Percent Similarity:	31.28%	Conservative:	152
Best Local Similarity:	19.12%	Mismatches:	410
Query Match:	1.80%	Indels:	449
DB:	3	Gaps:	58

US-10-664-859-1 (1-6909) x US-08-701-154A-5 (1-1061)

```
Qy       2952 TTGAAACCCCAATTCAATCAACATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGC 3011
```


Db 10 ValAsnProSerIleSerProAlaHisGlyValAlaArgSerSerThrLeuProProSer 29
Qy 3012 AACTCTTTTGTCGACCAGTCTGATCCTATGGGCAACGAAACTGAATTGATGTGCTGGGAA 3071
Db 30 AsnSerAsnMet---------- 33
Qy 3072 GGCGGATCCTCAAACACCAGTAGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATC 3131
Db 34 ---ValSerThrArgIleAsnArgGln---GlnSerSerAspLeuHisSerSerSerHis 51
Qy 3132 AGTACATCCAGCGAGTCACAG----------GCAATAAAGATACTGGAA 3170
Db 52 SerAsnSerSerAsnSerGlnGlySerPheGlyCysSerProGlySerGlnIleValAla 71
Qy 3171 GCAGCTGGCGTTGATTTGGGACAGGTC---ACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
Db 72 AsnValAlaLeuAsnLysGlyGlnAlaSerSerGlnSerSerLysProSerLeuAsnLeu 91
Qy 3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
Db 92 AsnAsnProProMetGluGlyThrGlyIleSerLeuAla---GlnPheMetSerPro--- 109
Qy 3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
Db 110 ---ArgArgGlnValThrSerGlyLeuAlaThrArgProArgMetProAsnAsnSerPhe 128
Qy 3348 CCTGAAAATGAGAATTCAGTAGGAGCT----------AAT 3377
Db 129 ProProAsnIleSerThrLeuSerSerProValGlyMetThrSerSerAlaCysAsnAsn 148
Qy 3378 GTAAGCTCACAGATAACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGA 3437
Db 149 AsnAsnArgSerTyrSerAsnIleProValThrSerLeuGlnGlyMetAsnGluGlyPro 168
Qy 3438 GGCGGATCTATT----------ATAAATCCGACGATGCGACAACTGCAT--------- 3476
Db 169 AsnAsnSerValGlyPheSerAlaSerSerProValLeuArgGlnMetSerSerGlnAsn 188
Qy 3477 ----------ATGCCAGGTAACGCCAAATCG---------- 3497
Db 189 SerProSerArgLeuAsnIleGlnProAlaLysAlaGluSerLysAspAsnLysGluIle 208
Qy 3498 ----------GAGCTCTTATCGGCGACAAGTTCAGGA----------CTT 3527
Db 209 AlaSerThrLeuAsnGluMetIleGlnSerAspAsnSerSerSerAspGlyLysProLeu 228
Qy 3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
Db 229 AspSerGlyLeuLeuHisAsnAsnAspArgLeuSerAsp----------GlyAsp 243
Qy 3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
Db 244 SerLysTyrSerGlnThrSerHisLysLeu---------ValGlnLeuLeuThr---Thr 259
Qy 3648 ACTGCA----------GCTGGAGTAAATGTCAATATGCATTGC----------TCA 3683
Db 260 ThrAlaGluGlnGlnLeuArgHisAlaAspIleAspThrSerCysLysAspValLeuSer 279
Qy 3684 AGCTCCGGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTT 3743
Db 280 CysThrGlyThrSerAsnSerAlaSerAlaAsnSerSerGly---------- 293
Qy 3744 GGCAACACAAGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAAT 3803
Db 294 ----------GlySerCysProSer----------Ser 299
Qy 3804 CAAGATAGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATT 3863
Db 300 HisSerSerLeuThrAlaArgHis----------LysIle 309
Qy 3864 CAACATCAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACT 3923
Db 310 LeuHisArgLeuLeuGlnGlu----------GlySerProSerAspIleThrThrLeu 325
Qy 3924 GTAGTACCACAACAGCAAACCCCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTG 3983
Db 326 SerValGluProAspLysLysAspSerAlaSerThrSerValSerValThrGlyGlnVal 345
Qy 3983 ---------- 3983
Db 346 GlnGlyAsnSerSerIleLysLeuGluLeuAspAlaSerLysLysLysGluSerLysAsp 365
Qy 3984 ----------CGACCCCTGCAAGGTCCACCT--- 4004
Db 366 HisGlnLeuLeuArgTyrLeuLeuAspLysAspGluLysAspLeuArgSerThrProAsn 385
Qy 4005 ----------CCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACT 4052
Db 386 LeuSerLeuAspAspValLysValLysValGluLysLysGluGlnMetAspProCysAsn 405
Qy 4053 CAATCGCCCAATCCC---------TCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGG 4103
Db 406 ThrAsnProThrProMetThrLysAlaThrProGluGluIleLysLeuGluAlaGlnSer 425
Qy 4104 ACAACCGCAGCAGTCATGGGA----------TTGCCGACC---------- 4133
Db 426 GlnPheThrAlaAspLeuAspGlnPheAspGlnLeuLeuProThrLeuGluLysAlaAla 445
Qy 4134 AACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAAT--- 4190
Db 446 GlnLeuProGlyLeuCysGluThrAspArgMetAspGlyAlaValThrSerValThrIle 465
Qy 4191 ----------ACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCAAAC 4235
Db 466 LysSerGluIleLeuProAlaSerLeuGlnSerAlaThrAlaArgProThrSerArgLeu 485
Qy 4236 AAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACC 4295
Db 486 AsnArgLeuProGluLeuGluLeuGluAlaIleAspAsnGlnPheGlyGlnProGlyThr 505
Qy 4296 GGA----------TCGTCAAGCGTTCTTACGCATAACTTAAGCAGCAACCCA 4337
Db 506 GlyAspGlnIleProTrpThrAsnAsnThrValThrAlaIleAsnGlnSerLysSerGlu 525
Qy 4338 ----------AGTACCCCCTTATCTCATCTA------TCCCCAAAGGAATTTGAGTCT 4379
Db 526 AspGlnCysIleSerSerGlnLeuAspGluLeuLeuCysProProThrThrValGluGly 545
Qy 4380 TTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTTTTT------AAAGACAAA 4433
Db 546 ArgAsnAspGluLysAlaLeuLeuGluGlnLeuValSerPheLeuSerGlyLysAspGlu 565
Qy 4434 TCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTC------GGAAATTTG 4487
Db 566 ThrGluLeuAlaGluLeuAspArgAlaLeuGlyIleAspLysLeuValGlnGlyGlyGly 585
Qy 4488 ATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCATTTCTTG 4547
Db 586 LeuAspValLeuSerGluArgProProGlnGlnAlaThrProProLeu--IleMetG 605
Qy 4548 AATCC------AATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAATTAAACACG 4601
Db 605 luGluArgProAsnLeuTyrSerGln----------ProTyrS 616
Qy 4602 AAAATTCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTT 4661
Db 616 erSerProPheProThrAla---------- 622
Qy 4662 ATTATAATATT---TCTACAGGTGATAACATGAAAAGTAGGCGACCAAGC---------- 4708
Db 623 -----AsnLeuProSerProPheGlnGlyMetValArgGlnLysProSerLeuGlyThrM 641
Qy 4709 ----------CCACAGGGTCAGCGGTCACCA---------- 4729
Db 641 etProValGlnValThrProProArgGlyAlaPheSerProGlyMetGlyMetGlnProA 661



Search completed: November 2, 2005, 22:23:59
Job time : 275 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 22:16:48 ; Search time 821.5 Seconds
(without alignments)
6505.484 Million cell updates/sec

Title: US-10-664-859-1
Perfect score: 12037
Sequence: 1 acgagtgcttctcttattat.....aaaaaaaaaaaaaaaaaaaa 6909

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10664859/runat_02112005_171806_3593/app_query.fasta_1.7047
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10664859_@CGN_1_1_995_@runat_02112005_171806_3593 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7257	60.3	1464	5	AAB71228	Aab71228 D. melano
2	7257	60.3	1464	7	ABW01527	Abw01527 Drosophil
3	7257	60.3	1464	8	ADJ71911	Adj71911 Fruit fly
4	7040	58.5	1429	4	ABB58779	Abb58779 Drosophil
5	570	4.7	112	7	ABW01532	Abw01532 Drosophil
6	570	4.7	112	8	ADJ71898	Adj71898 Fruit fly
7	321.5	2.7	1435	4	ABB11808	Abb11808 Human BCL
8	314	2.6	1394	8	ADQ18945	Adq18945 Human sof
9	312	2.6	1426	5	AAB71229	Aab71229 Human leg
10	312	2.6	1426	7	ABW01534	Abw01534 Human lgs
11	312	2.6	1426	7	ADJ70152	Adj70152 Human hea
12	312	2.6	1426	8	ADJ71903	Adj71903 Human Lgs
13	296.5	2.5	1494	5	AAU78460	Aau78460 Mouse bet
14	270.5	2.2	1596	6	AAE36095	Aae36095 Drosophil
15	269.5	2.2	1594	4	ABB63688	Abb63688 Drosophil
16	258.5	2.1	1366	4	ABB66921	Abb66921 Drosophil
17	258.5	2.1	2151	4	ABB60086	Abb60086 Drosophil
18	247.5	2.1	1902	6	ABJ38695	Abj38695 Human nuc
19	242.5	2.0	1161	4	ABB70667	Abb70667 Drosophil
20	241.5	2.0	1778	4	ABB60081	Abb60081 Drosophil
21	241	2.0	2280	4	ABB61650	Abb61650 Drosophil
22	240	2.0	2112	4	ABB60403	Abb60403 Drosophil
23	233.5	1.9	1243	7	ADJ70158	Adj70158 Human hea
24	233.5	1.9	1441	7	ADC23487	Adc23487 Receptor
25	233	1.9	2703	4	ABB63299	Abb63299 Drosophil
26	233	1.9	2703	5	ABG70019	Abg70019 Larval vi
27	233	1.9	3190	4	AAB84634	Aab84634 Amino aci
28	231.5	1.9	1844	3	AAB18250	Aab18250 Plasmodiu
29	230.5	1.9	3275	4	ABB70437	Abb70437 Drosophil
30	230.5	1.9	3275	7	ADK11301	Adk11301 Drosophil
31	229.5	1.9	2005	6	ABG74435	Abg74435 Human nuc
32	228.5	1.9	2063	4	ABG17147	Abg17147 Novel hum
33	228.5	1.9	2063	6	ABG74436	Abg74436 Human nuc
34	228.5	1.9	2063	8	ADQ91924	Adq91924 Human act
35	228	1.9	1424	6	ABU03694	Abu03694 Human exp
36	228	1.9	1424	6	ABU03681	Abu03681 Human exp
37	228	1.9	1424	6	ABU03684	Abu03684 Human exp
38	228	1.9	1424	7	ADE76188	Ade76188 Human AIB
39	225.5	1.9	2063	4	AAM78710	Aam78710 Human pro
40	225.5	1.9	2063	4	AAM40064	Aam40064 Human pol
41	223.5	1.9	1522	2	AAY21975	Aay21975 Human ste
42	223.5	1.9	1522	6	ABU03689	Abu03689 Human exp
43	222.5	1.8	1063	5	ABB54168	Abb54168 Lactococc
44	222.5	1.8	1441	4	AAE12570	Aae12570 Human ste
45	222.5	1.8	1441	5	AAE25569	Aae25569 Human ste

ALIGNMENTS

RESULT 1
AAB71228
ID AAB71228 standard; protein; 1464 AA.
XX
AC AAB71228;
XX
DT 18-NOV-2002 (first entry)
XX
DE D. melanogaster lgs protein.
XX
KW Legless; fruitfly; lgs; Wnt/Wingless signaling pathway; Wnt; Wg;
KW tissue proliferation; tumour; cytostatic; cellular disorder; colon;
KW blood disorder; cancer; breast; head and neck cancer; brain; thyroid;
KW medulloblastoma; skin cancer; tissue regeneration; tissue repair.
XX
OS Drosophila melanogaster.
XX
PN US2002086986-A1.
XX
PD 04-JUL-2002.
XX
PF 27-JUL-2001; 2001US-00915543.
XX
PR 28-JUL-2000; 2000US-0221502P.
XX
PA (BASL/) BASLER K.
PA (BRUN/) BRUNNER E.
PA (FROE/) FROESCH B.
PA (KRAM/) KRAMPS T.
PA (PETE/) PETER O.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2002-635689/68.

```
DR   N-PSDB; AAF88466.
XX
PT   Novel polypeptide useful in therapeutic method for treating disorders of
PT   cell fate such as cell differentiation or cell proliferation.
XX
PS   Example II; Fig 2; 41pp; English.
XX
CC   This invention describes a novel polypeptide sharing one or more
CC   homologous amino acid domains with the legless (lgs) protein, a
CC   downstream component of the Wnt/Wingless (Wnt/Wg) signaling pathway
CC   involved in the formation and maintenance of spatial arrangements and
CC   proliferation of tissues during development, and in the formation and
CC   growth of many human tumours. The products of the invention have
CC   cytostatic activity and can be used to treat cellular disorders, blood
CC   disorders and cancers caused by over-stimulation of the Wnt pathway,
CC   where the cancerous condition is colon, breast, head and neck, brain,
CC   thyroid, medulloblastoma or skin cancer. The product could also be used
CC   to promote tissue regeneration and repair. This sequence represents the
CC   Drosophila melanogaster (fruitfly) legless (lgs) protein described in the
CC   disclosure of the invention
XX
SQ   Sequence 1464 AA;

Alignment Scores:
Pred. No.:                0          Length:          1464
Score:                    7257.00    Matches:         1460
Percent Similarity:       73.01%     Conservative:    1
Best Local Similarity:    72.96%     Mismatches:      3
Query Match:              60.29%     Indels:          537
DB:                       5          Gaps:            5

US-10-664-859-1 (1-6909) x AAB71228 (1-1464)

Qy        468 ATGCCCCGCAGTCCAACCCAACAACAGCCGCAACCAAACTCCGATGCCTCCTCAACAAGT 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetProArgSerProThrGlnGlnGlnProGlnProAsnSerAspAlaSerSerThrSer 20

Qy        528 GCATCTGGATCAAATCCTGGAGCAGCGATCGGAAATGGGGACTCGGCGGCGAGCAGAAGT 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSerAlaAlaSerArgSer 40

Qy        588 TCTCCGAAGACCCTTAATAGCGAACCCTTTTCTACTTTGTCGCCGGGTAAGACTTGTATT 647
              |||||||||||||||||||||||||||||||||||
Db         41 SerProLysThrLeuAsnSerGluProPheSerThrLeu-Ser---------------- 54

Qy        648 GATTTCTCTTTGTCCGGAATTATAACAACTTTCTGTGTTTCCAGATCAAATAAAATTGAC 707
                                                  ||||||||||||||||||||
Db         55 ----------------------------------------ProAspGlnIleLysLeuTh 61

Qy        708 GCCAGAAGAAGGCACTGAGAAAAGCGGACTATCAACTAGTGATAAAGCTGCCACTGGAGG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 rProGluGluGlyThrGluLysSerGlyLeuSerThrSerAspLysAlaAlaThrGlyGl 81

Qy        768 AGCCCCAGGCAGTGGAAATAATCTGCCCGAGGGACAAACTATGCTAAGGCAGAACTCTAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 yAlaProGlySerGlyAsnAsnLeuProGluGlyGlnThrMetLeuArgGlnAsnSerTh 101

Qy        828 GAGCACAATCAACTCGTGCCTAGTCGCTTCTCCACAAAACTCCAGTGAACACTCGAATAG 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 rSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSerGluHisSerAsnSe 121

Qy        888 CAGCAATGTGTCTGCTACAGTGGGCCTTACTCAGATGGTAGATTGTGACGAGCAATCGAA 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 rSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCysAspGluGlnSerLy 141

Qy        948 GAAAAACAAATGTAGTGTGAAGGACGAGGAAGCTGGTAAGACTGCCCTACAAATGGTTTA 1007
              ||||||||||||||||||||||||||||||||||
Db        141 sLysAsnLysCysSerValLysAspGluGluAla------------------------- 152

Qy       1008 AAATTTTAAAATGTATTGGCGTTCACCTTTGTTAATCATTTAATTGTTTTTTTTTTGCTA 1067

Db        152 ------------------------------------------------------------ 152
```

```
Qy       1068 TACTTACAATTTTAGTTTTAAACTTGTAAACTTGACTAAAACTCGCGAAGCTCGGATCAA 1127

Db        152 ------------------------------------------------------------ 152

Qy       1128 AACAGACATTTTCTTGGAACCGTAATTAAGCTCATAAAAATATTAATTCATCTTGATGGA 1187

Db        152 ------------------------------------------------------------ 152

Qy       1188 ATGCATATCATAGATGTACTCAAACATCTCAAGAAAGACCTCAAATTGGATCAACTAATT 1247

Db        152 ------------------------------------------------------------ 152

Qy       1248 AGTTTGAGAAAAAATTGCTGTACTTTTAAGAATATATTAATTTAAAAATTTGCTGAGTGA 1307

Db        152 ------------------------------------------------------------ 152

Qy       1308 AATGATATAATAGTCACAATAATTTTTAGTTAAACTGCTAAAGCATTTTGAATAGCCGTG 1367

Db        152 ------------------------------------------------------------ 152

Qy       1368 CTACGCAGATGCTACTAGACGCGGTGTAAAAGCTAATTTTTATTTAAAAGCTGTCCTAAT 1427

Db        152 ------------------------------------------------------------ 152

Qy       1428 ATTCCATAACCATTAATGTCCCATTTCAGAAATAAGTTCTAATAAAGCAAAAGGTCAAGC 1487
                                              ||||||||||||||||||||||||||||
Db        153 ----------------------------GluIleSerSerAsnLysAlaLysGlyGlnAl 163

Qy       1488 AGCTGGTGGCGGCTGCGAAACAGGTTCTACATCCAGTTTGACTGTCAAGGAAGAACCCAC 1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 aAlaGlyGlyGlyCysGluThrGlySerThrSerSerLeuThrValLysGluGluProTh 183

Qy       1548 CGATGTCTTAGGCAGTTTAGTAAATATGAAAAAAGAAGAAAGAGAAAATCATTCGCCAAC 1607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 rAspValLeuGlySerLeuValAsnMetLysLysGluGluArgGluAsnHisSerProTh 203

Qy       1608 GATGTCCCCTGTTGGTTTTGGTTCAATTGGTAATGCACAGGACAACTCCGCTACACCGGG 1667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 rMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAsnSerAlaThr----- 221

Qy       1668 TAAGTTTTAAGAGATCCATATAAAGCAAATAACAAGAATTAATGTCAGTTACCAATTTTA 1727

Db        221 ------------------------------------------------------------ 221

Qy       1728 TTTGATAGTCAAAGAACTACTATAGCGATATCTCCTGCCTTTTAATTTTATTTTAATTAG 1787

Db        221 ------------------------------------------------------------ 221

Qy       1788 GAAATACGAATATTTCTAATTTGTAAAATAAAATTGATTAATTAACTAGAATTTAAAAAC 1847

Db        221 ------------------------------------------------------------ 221

Qy       1848 CTTTTGAATTAGGACATACCCTTCCAAAAATCAGTAATCATTGGGAACGAGAGTGTGGTC 1907

Db        221 ------------------------------------------------------------ 221

Qy       1908 CCGAAGGAGACTACTATAAAACCTTTTGAGCTATCTGATACTGCACGCTACTAAAAATGA 1967

Db        221 ------------------------------------------------------------ 221

Qy       1968 TTAGTTTAGGAAAATGGGTGTAATTTTGTAGGAAGTTTTCATTTTAGAAGAAATGTGATT 2027

Db        221 ------------------------------------------------------------ 221

Qy       2028 ATTTTATTAAACCCCTTCAAGCGGAACTACATTTGTTCTACGATATTTTGGAAAAACAAA 2087

Db        221 ------------------------------------------------------------ 221

Qy       2088 TGGTTAAGTTGGAAAGTGCCTATAAAACAGAATTCCACGGTTTCAAATACTAACCAGGTT 2147

Db        221 ------------------------------------------------------------ 221

Qy       2148 TTTGATTTAATTTTGATTAAATGAGAAATTATCACACTTCAGTTAAAATGTTTAATTCGA 2207
```

```
Db        221 ------------------------------------------------------------ 221
Qy       2208 TTAAGGTCGGACAATCACAGCAGATTTCCATTTTTGCGTGTATATATAGAAGTCGCCTTC 2267
Db        221 ------------------------------------------------------------ 221
Qy       2268 ACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCCCGCAGTGATTTATATAGAT 2327
Db        221 ------------------------------------------------------------ 221
Qy       2328 GATTTACGAGTTATTTAATTTTTTATGGTGTATTTTAATAAATATCTTATTTATTCATTT 2387
Db        221 ------------------------------------------------------------ 221
Qy       2388 TACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCC 2447
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 -Pro--ValLysIleGluArgIleSerAsnAspSerThrThrGluLysLysGlySerSer 240
Qy       2448 TTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTT 2507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGlnLeuAsnProAspPhe 260
Qy       2508 ATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCA 2567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuValSerGlyValGlyPro 280
Qy       2568 ATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGA 2627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 IleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThrAlaAsnAlaAsnGly 300
Qy       2628 ATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCACATATTCGTG 2687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGlnAsnHisIlePheVal 320
Qy       2688 TTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTTCAAACT 2747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 PheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSerGlyGlnPheGlnThr 340
Qy       2748 ATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTTTTTATG 2807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeuGluAspPhePheMet 360
Qy       2808 AAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATA 2867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerValGlyMetProTrpIle 380
Qy       2868 GGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAG 2927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLysIleThrGlnGlnGln 400
Qy       2928 CCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAACAGCAAA 2987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGlnHisGluAsnSerLys 420
Qy       2988 CGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGATCCTATGGGCAAC 3047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgSerThrValSerAlaProSerAsnSerPheValAspGlnSerAspProMetGlyAsn 440
Qy       3048 GAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAAC 3107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GluThrGluLeuMetCysTrpGluGlyGlySerSerAsnThrSerArgSerGlyGlnAsn 460
Qy       3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerArgAsnHisValAspSerIleSerThrSerSerGluSerGlnAlaIleLysIleLeu 480
Qy       3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAspProGlyLeuThrThr 500
Qy       3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        501 GluAsnAsnIleValSerLeuGlnGlyValLysValProAspGluAsnLeuThrProGln 520
Qy       3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnArgGlnHisArgGluGluGlnLeuAlaLysIleLysLysMetAsnGlnPheLeuPhe 540
Qy       3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIleThrLysIleProGly 560
Qy       3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 AspLeuMetMetGlyMetSerGlyGlyGlyGlyGlySerIleIleAsnProThrMetArg 580
Qy       3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAlaThrSerSerGlyLeu 600
Qy       3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 SerGluAspValMetHisProGlyAspValIleSerAspMetGlyAlaValIleGlyCys 620
Qy       3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 AsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGlyValValThrGlyThr 640
Qy       3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGlyAlaProAsnGlyAsn 660
Qy       3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 MetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThrSerCysAsnValIle 680
Qy       3768 GGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACCCATTCACAT 3827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 GlyThrAlaProAspMetSerLysGluValLeuAsnGlnAspSerArgThrHisSerHis 700
Qy       3828 CAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAATTTTTCGAAGAACGC 3887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 GlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHisGlnPhePheGluGluArg 720
Qy       3888 CTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACCCCT 3947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LeuLysGlyGlyLysProArgGlnValThrGlyThrValValProGlnGlnGlnThrPro 740
Qy       3948 TCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCT 4007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 SerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeuGlnGlyProProPro 760
Qy       4008 CCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCC 4067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 ProTyrHisSerIleGlnArgSerAlaSerValProIleAlaThrGlnSerProAsnPro 780
Qy       4068 TCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTG 4127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SerSerProAsnAsnLeuSerLeuProSerProArgThrThrAlaAlaValMetGlyLeu 800
Qy       4128 CCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCT 4187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 ProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGlySerValProGlnAla 820
Qy       4188 AATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTT 4247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 AsnThrSerThrValGlnAlaGlyThrThrThrValLeuSerAlaAsnLysAsnCysPhe 840
Qy       4248 CAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGC 4307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsnThrGlySerSerSer 860
Qy       4308 GTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG 4367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 ValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSerHisLeuSerProLys 880
```

```
Qy       4368 GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTTTTTAAA 4427
             ||||||||||||||||||||||||
Db        881 GluPheGluSerPheGlyGlnSerSer---------------------------------- 889

Qy       4428 GACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGAAATTTG 4487
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 ------------------------------------------------------------ 889

Qy       4488 ATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCATTTCTTG 4547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 ------------------------------------------------------------ 889

Qy       4548 AATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAATTAAACACGAAAATT 4607
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 ------------------------------------------------------------ 889

Qy       4608 CATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATA 4667
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 ------------------------------------------------------------ 889

Qy       4668 ATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCAC 4727
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        890 ------Ala--GlyAspAsnMetLysSerArgArgProSerProGlnGlyGlnArgSerP 907

Qy       4728 CAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTT 4787
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 roValAsnSerLeuIleGluAlaAsnLysAspValArgPheAlaAlaSerSerProGlyP 927

Qy       4788 TTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGG 4847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 heAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuAsnAlaTyrLysMetG 947

Qy       4848 GCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTT 4907
             |||||||||||||||||||||  :::
Db        947 lySerThrAsnIleGlnMetGlu--Arg---------------------------------- 955

Qy       4908 AATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCG 4967
                                  ||||||||||||||||||||||||||||||||||||||
Db        956 ----------------------GlnAlaSerAlaGlnGlyGlySerValGlnPheSerAr 968

Qy       4968 GCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCCACCAAACAAGAT 5027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 gArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgProProProAsnLysMe 988

Qy       5028 GACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTTG 5087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 tThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGlnGlnLeuThrSerCy 1008

Qy       5088 CGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCGGG 5147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 sValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMetMetGlyGlyProGl 1028

Qy       5148 ACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGATGGATCAGGAATAGA 5207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 yProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAspGlySerGlyIleAs 1048

Qy       5208 TACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAACTCAATGGGTCCCCG 5267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 pThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAsnSerMetGlyProAr 1068

Qy       5268 AATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACCGCCTGGCTTTAATCC 5327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 gMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyProProGlyPheAsnPr 1088

Qy       5328 TAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAA 5387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 oAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerGlyCysGlySerAlaAs 1108

Qy       5388 CTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGT 5447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 nSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMetGlyArgMetProVa 1128

Qy       5448 CAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAGTACCCCAAACACCAT 5507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 lAsnPheGlySerAsnPheAsnProAsnIleGlnValLysAlaSerThrProAsnThrIl 1148

Qy       5508 ACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGGAGCTAATAA 5567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1148 eGlnTyrMetProValArgAlaGlnAsnAlaAsnAsnAsnAsnAsnGlyAlaAsnAs 1168

Qy       5568 TGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGC 5627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 nValArgMetProProSerLeuGluPheLeuGlnArgTyrAlaAsnProGlnMetGlyAl 1188

Qy       5628 TGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCC 5687
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 aValGlyAsnGlySerProIleCysProProSerAlaSerAspGlyThrProGlyMetPr 1208

Qy       5688 AGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTCTTCCGGAGAGCAACA 5747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 oGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSerSerGlyGluGlnHi 1228

Qy       5748 CCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTG 5807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 sGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAsnPhePheGlnAsnCy 1248

Qy       5808 CAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTACCCGGCCATGACGGATCAATGAA 5867
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 sAsnGlnMetSerIleValAspGluGluGlyGlyLeuProGlyHisAspGlySerMetAs 1268

Qy       5868 TATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCATGCGGCCAAATGTAAT 5927
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 nIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMetArgProAsnValMe 1288

Qy       5928 GGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACAGTCATCGGATGGTAT 5987
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 tGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGlnSerSerAspGlyIl 1308

Qy       5988 TGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTGCAACAGCGCTGGACC 6047
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 eAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCysAsnSerAlaGlyPr 1328

Qy       6048 ACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACAACACATAAAGAACAT 6107
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 oHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGlnHisIleLysAsnIl 1348

Qy       6108 ACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACAAGGGCAGATCCAACT 6167
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 eProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGlnGlyGlnIleGlnLe 1368

Qy       6168 GCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTACTAATAATAATTTAAT 6227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 uHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProThrAsnAsnAsnLeuMe 1388

Qy       6228 GTCAACTGCCGGAAGTGTCAGTGCTACTAACGGTGTCTCTGGCATCAATTTCGTAGGTCC 6287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 tSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIleAsnPheValGlyPr 1408

Qy       6288 CTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCAGCAGCAGTTATATGC 6347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 oSerSerThrAspLeuLysTyrAlaGlnGlnTyrHisSerPheGlnGlnGlnLeuTyrAl 1428

Qy       6348 TACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCACCAGAGCAACATGAT 6407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1428 aThrAsnThrArgSerGlnGlnGlnGlnHisMetHisGlnGlnHisGlnSerAsnMetIl 1448

Qy       6408 AACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTCAACAAA 6456
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 eThrMetProProAsnLeuSerProAsnProThrPhePheValAsnLys 1464
```

```
RESULT 2
ABW01527
ID   ABW01527 standard; protein; 1464 AA.
XX
AC   ABW01527;
```

XX
DT 15-JAN-2004 (first entry)
XX
DE Drosophila species legless (lgs) protein.
XX
KW Legless protein; lgs; cell fate disorder; blood disease; gene therapy;
KW cancer; tissue regeneration; tissue repair; cytostatic.
XX
OS Drosophila sp.
XX
PN US2003114413-A1.
XX
PD 19-JUN-2003.
XX
PF 19-DEC-2002; 2002US-00322579.
XX
PR 28-JUL-2000; 2000US-0221502P.
PR 27-JUL-2001; 2001US-00915543.
XX
PA (UYZU-) UNIV ZURICH.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2003-829432/77.
DR N-PSDB; AAD62641.
XX
PT Novel Lgs polypeptide useful for isolation of Lgs-binding proteins,
PT diagnosing disorders of cell fate, treating diseases such as cancer.
XX
PS Claim 5; Fig 2; 0pp; English.
XX
CC The invention relates to novel legless (lgs) proteins and polynucleotides
CC encoding such proteins. Lgs sequences are useful for the treatment of
CC disorders of cell fate such as differentiation or proliferation. The
CC invention is used to treat blood disease or a cancerous condition
CC characterised by over-stimulation of the Wnt pathway such as colon,
CC breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and
CC is administered to prevent progression from a pre-neoplastic or non-
CC malignant condition to a neoplastic or malignant state. It is
CC administered to promote tissue regeneration and repair. The invention is
CC also useful in the therapy of diseases cost by an over-activation of Wg
CC pathway. It is useful for reducing lgs gene expression in an invertebrate
CC or vertebrate organism or an invertebrate or vertebrate cell line. The
CC invention is also useful in gene therapy. The present sequence is
CC Drosophila species legless (lgs) protein
XX
SQ Sequence 1464 AA;

Alignment Scores:

Pred. No.:	0	Length:	1464
Score:	7257.00	Matches:	1460
Percent Similarity:	73.01%	Conservative:	1
Best Local Similarity:	72.96%	Mismatches:	3
Query Match:	60.29%	Indels:	537
DB:	7	Gaps:	5

US-10-664-859-1 (1-6909) x ABW01527 (1-1464)

```
Qy        468 ATGCCCCGCAGTCCAACCCAACAACAGCCGCAACCAAACTCCGATGCCTCCTCAACAAGT 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetProArgSerProThrGlnGlnGlnProGlnProAsnSerAspAlaSerSerThrSer 20

Qy        528 GCATCTGGATCAAATCCTGGAGCAGCGATCGGAAATGGGGACTCGGCGGCGAGCAGAAGT 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSerAlaAlaSerArgSer 40

Qy        588 TCTCCGAAGACCCTTAATAGCGAACCCTTTTCTACTTTGTCGCCGGGTAAGACTTGTATT 647
              ||||||||||||||||||||||||||||||||||
Db         41 SerProLysThrLeuAsnSerGluProPheSerThrLeu-Ser---------------- 54

Qy        648 GATTTCTCTTTGTCCGGAATTATAACAACTTTCTGTGTTTCCAGATCAAATAAAATTGAC 707
                                                  ||||||||||||||||||||
Db         55 ----------------------------------------ProAspGlnIleLysLeuTh 61

Qy        708 GCCAGAAGAAGGCACTGAGAAAAGCGGACTATCAACTAGTGATAAAGCTGCCACTGGAGG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 rProGluGluGlyThrGluLysSerGlyLeuSerThrSerAspLysAlaAlaThrGlyGl 81

Qy        768 AGCCCCAGGCAGTGGAAATAATCTGCCCGAGGGACAAACTATGCTAAGGCAGAACTCTAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 yAlaProGlySerGlyAsnAsnLeuProGluGlyGlnThrMetLeuArgGlnAsnSerTh 101

Qy        828 GAGCACAATCAACTCGTGCCTAGTCGCTTCTCCACAAAACTCCAGTGAACACTCGAATAG 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 rSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSerGluHisSerAsnSe 121

Qy        888 CAGCAATGTGTCTGCTACAGTGGGCCTTACTCAGATGGTAGATTGTGACGAGCAATCGAA 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 rSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCysAspGluGlnSerLy 141

Qy        948 GAAAAACAAATGTAGTGTGAAGGACGAGGAAGCTGGTAAGACTGCCCTACAAATGGTTTA 1007
              ||||||||||||||||||||||||||||||
Db        141 sLysAsnLysCysSerValLysAspGluGluAla---------------------------- 152

Qy       1008 AAATTTTAAAATGTATTGGCGTTCACCTTTGTTAATCATTTAATTGTTTTTTTTTTGCTA 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1068 TACTTACAATTTTAGTTTTAAACTTGTAAACTTGACTAAAACTCGCGAAGCTCGGATCAA 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1128 AACAGACATTTTCTTGGAACCGTAATTAAGCTCATAAAAATATTAATTCATCTTGATGGA 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1188 ATGCATATCATAGATGTACTCAAACATCTCAAGAAAGACCTCAAATTGGATCAACTAATT 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1248 AGTTTGAGAAAAAATTGCTGTACTTTTAAGAATATATTAATTTAAAAATTTGCTGAGTGA 1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1308 AATGATATAATAGTCACAATAATTTTTAGTTAAACTGCTAAAGCATTTTGAATAGCCGTG 1367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1368 CTACGCAGATGCTACTAGACGCGGTGTAAAAGCTAATTTTTATTTAAAAGCTGTCCTAAT 1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 ---------------------------------------------------------- 152

Qy       1428 ATTCCATAACCATTAATGTCCCATTTCAGAAATAAGTTCTAATAAAGCAAAAGGTCAAGC 1487
                                                  ||||||||||||||||||||||||||||||
Db        153 ---------------------------GluIleSerSerAsnLysAlaLysGlyGlnAl 163

Qy       1488 AGCTGGTGGCGGCTGCGAAACAGGTTCTACATCCAGTTTGACTGTCAAGGAAGAACCCAC 1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 aAlaGlyGlyGlyCysGluThrGlySerThrSerSerLeuThrValLysGluGluProTh 183

Qy       1548 CGATGTCTTAGGCAGTTTAGTAAATATGAAAAAAGAAGAAAGAGAAAATCATTCGCCAAC 1607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 rAspValLeuGlySerLeuValAsnMetLysLysGluGluArgGluAsnHisSerProTh 203

Qy       1608 GATGTCCCCTGTTGGTTTTGGTTCAATTGGTAATGCACAGGACAACTCCGCTACACCGGG 1667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 rMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAsnSerAlaThr----- 221

Qy       1668 TAAGTTTTAAGAGATCCATATAAAGCAAATAACAAGAATTAATGTCAGTTACCAATTTTA 1727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 ---------------------------------------------------------- 221

Qy       1728 TTTGATAGTCAAAGAACTACTATAGCGATATCTCCTGCCTTTTAATTTTATTTTAATTAG 1787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 ---------------------------------------------------------- 221
```

```
Qy       1788 GAAATACGAATATTTCTAATTTGTAAAATAAAATTGATTAATTAACTAGAATTTAAAAAC 1847
Db        221 ------------------------------------------------------------ 221
Qy       1848 CTTTTGAATTAGGACATACCCTTCCAAAAATCAGTAATCATTGGGAACGAGAGTGTGGTC 1907
Db        221 ------------------------------------------------------------ 221
Qy       1908 CCGAAGGAGACTACTATAAAACCTTTTGAGCTATCTGATACTGCACGCTACTAAAAATGA 1967
Db        221 ------------------------------------------------------------ 221
Qy       1968 TTAGTTTAGGAAAATGGGTGTAATTTTGTAGGAAGTTTTCATTTTAGAAGAAATGTGATT 2027
Db        221 ------------------------------------------------------------ 221
Qy       2028 ATTTTATTAAACCCCTTCAAGCGGAACTACATTTGTTCTACGATATTTTGGAAAAACAAA 2087
Db        221 ------------------------------------------------------------ 221
Qy       2088 TGGTTAAGTTGGAAAGTGCCTATAAAACAGAATTCCACGGTTTCAAATACTAACCAGGTT 2147
Db        221 ------------------------------------------------------------ 221
Qy       2148 TTTGATTTAATTTTGATTAAATGAGAAATTATCACACTTCAGTTAAAATGTTTAATTCGA 2207
Db        221 ------------------------------------------------------------ 221
Qy       2208 TTAAGGTCGGACAATCACAGCAGATTTCCATTTTTGCGTGTATATATAGAAGTCGCCTTC 2267
Db        221 ------------------------------------------------------------ 221
Qy       2268 ACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCCCGCAGTGATTTATATAGAT 2327
Db        221 ------------------------------------------------------------ 221
Qy       2328 GATTTACGAGTTATTTAATTTTTTATGGTGTATTTTAATAAATATCTTATTTATTCATTT 2387
Db        221 ------------------------------------------------------------ 221
Qy       2388 TACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCC 2447
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 -Pro--ValLysIleGluArgIleSerAsnAspSerThrThrGluLysLysGlySerSer 240
Qy       2448 TTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTT 2507
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 LeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGlnLeuAsnProAspPhe 260
Qy       2508 ATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCA 2567
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 IleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuValSerGlyValGlyPro 280
Qy       2568 ATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGA 2627
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 IleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThrAlaAsnAlaAsnGly 300
Qy       2628 ATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCACATATTCGTG 2687
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGlnAsnHisIlePheVal 320
Qy       2688 TTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTTCAAACT 2747
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 PheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSerGlyGlnPheGlnThr 340
Qy       2748 ATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTTTTTATG 2807
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 IleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeuGluAspPhePheMet 360
Qy       2808 AAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATA 2867
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerValGlyMetProTrpIle 380
Qy       2868 GGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAG 2927
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 GlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLysIleThrGlnGlnGln 400
Qy       2928 CCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAACAGCAAA 2987
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 ProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGlnHisGluAsnSerLys 420
Qy       2988 CGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGATCCTATGGGCAAC 3047
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgSerThrValSerAlaProSerAsnSerPheValAspGlnSerAspProMetGlyAsn 440
Qy       3048 GAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAAC 3107
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GluThrGluLeuMetCysTrpGluGlyGlySerSerAsnThrSerArgSerGlyGlnAsn 460
Qy       3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 SerArgAsnHisValAspSerIleSerThrSerSerGluSerGlnAlaIleLysIleLeu 480
Qy       3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GluAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAspProGlyLeuThrThr 500
Qy       3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GluAsnAsnIleValSerLeuGlnGlyValLysValProAspGluAsnLeuThrProGln 520
Qy       3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAAATGAATCAATTTCTTTTT 3347
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 GlnArgGlnHisArgGluGluGlnLeuAlaLysIleLysLysMetAsnGlnPheLeuPhe 540
Qy       3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIleThrLysIleProGly 560
Qy       3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 AspLeuMetMetGlyMetSerGlyGlyGlyGlyGlySerIleIleAsnProThrMetArg 580
Qy       3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 GlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAlaThrSerSerGlyLeu 600
Qy       3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 SerGluAspValMetHisProGlyAspValIleSerAspMetGlyAlaValIleGlyCys 620
Qy       3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 AsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGlyValValThrGlyThr 640
Qy       3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        641 ThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGlyAlaProAsnGlyAsn 660
Qy       3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 MetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThrSerCysAsnValIle 680
Qy       3768 GGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACCCATTCACAT 3827
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 GlyThrAlaProAspMetSerLysGluValLeuAsnGlnAspSerArgThrHisSerHis 700
Qy       3828 CAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAATTTTTCGAAGAACGC 3887
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 GlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHisGlnPhePheGluGluArg 720
Qy       3888 CTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACCCCT 3947
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LeuLysGlyGlyLysProArgGlnValThrGlyThrValValProGlnGlnGlnThrPro 740
Qy       3948 TCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCT 4007
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        741 SerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeuGlnGlyProProPro 760

Qy       4008 CCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCC 4067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 ProTyrHisSerIleGlnArgSerAlaSerValProIleAlaThrGlnSerProAsnPro 780

Qy       4068 TCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTG 4127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SerSerProAsnAsnLeuSerLeuProSerProArgThrThrAlaAlaValMetGlyLeu 800

Qy       4128 CCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCT 4187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 ProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGlySerValProGlnAla 820

Qy       4188 AATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTT 4247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 AsnThrSerThrValGlnAlaGlyThrThrThrValLeuSerAlaAsnLysAsnCysPhe 840

Qy       4248 CAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGC 4307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsnThrGlySerSerSer 860

Qy       4308 GTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG 4367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 ValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSerHisLeuSerProLys 880

Qy       4368 GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTTTTTAAA 4427
              ||||||||||||||||||||||||
Db        881 GluPheGluSerPheGlyGlnSerSer---------------------------------- 889

Qy       4428 GACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGAAATTTG 4487

Db        889 ---------------------------------------------------------- 889

Qy       4488 ATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCATTTCTTG 4547

Db        889 ---------------------------------------------------------- 889

Qy       4548 AATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAATTAAACACGAAAATT 4607

Db        889 ---------------------------------------------------------- 889

Qy       4608 CATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATA 4667

Db        889 ---------------------------------------------------------- 889

Qy       4668 ATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCAC 4727
                        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        890 ------Ala--GlyAspAsnMetLysSerArgArgProSerProGlnGlyGlnArgSerP 907

Qy       4728 CAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTT 4787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 roValAsnSerLeuIleGluAlaAsnLysAspValArgPheAlaAlaSerSerProGlyP 927

Qy       4788 TTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGG 4847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 heAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuAsnAlaTyrLysMetG 947

Qy       4848 GCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTT 4907
              ||||||||||||||||||||||  :::
Db        947 lySerThrAsnIleGlnMetGlu--Arg------------------------------- 955

Qy       4908 AATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCG 4967
                                    ||||||||||||||||||||||||||||||||||||
Db        956 ----------------------GlnAlaSerAlaGlnGlyGlySerValGlnPheSerAr 968

Qy       4968 GCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCCACCAAACAAGAT 5027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        968 gArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgProProProAsnLysMe 988

Qy       5028 GACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTTG 5087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 tThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGlnGlnLeuThrSerCy 1008
```

```
Qy       5088 CGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCGGG 5147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 sValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMetMetGlyGlyProGl 1028

Qy       5148 ACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGATGGATCAGGAATAGA 5207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 yProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAspGlySerGlyIleAs 1048

Qy       5208 TACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAACTCAATGGGTCCCCG 5267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 pThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAsnSerMetGlyProAr 1068

Qy       5268 AATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACCGCCTGGCTTTAATCC 5327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 gMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyProProGlyPheAsnPr 1088

Qy       5328 TAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAA 5387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 oAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerGlyCysGlySerAlaAs 1108

Qy       5388 CTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGT 5447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 nSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMetGlyArgMetProVa 1128

Qy       5448 CAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAGTACCCCAAACACCAT 5507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 lAsnPheGlySerAsnPheAsnProAsnIleGlnValLysAlaSerThrProAsnThrIl 1148

Qy       5508 ACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGGAGCTAATAA 5567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1148 eGlnTyrMetProValArgAlaGlnAsnAlaAsnAsnAsnAsnAsnGlyAlaAsnAs 1168

Qy       5568 TGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGC 5627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 nValArgMetProProSerLeuGluPheLeuGlnArgTyrAlaAsnProGlnMetGlyAl 1188

Qy       5628 TGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCC 5687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 aValGlyAsnGlySerProIleCysProProSerAlaSerAspGlyThrProGlyMetPr 1208

Qy       5688 AGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTCTTCCGGAGAGCAACA 5747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 oGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSerSerGlyGluGlnHi 1228

Qy       5748 CCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTG 5807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 sGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAsnPhePheGlnAsnCy 1248

Qy       5808 CAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTACCCGGCCATGACGGATCAATGAA 5867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 sAsnGlnMetSerIleValAspGluGluGlyGlyLeuProGlyHisAspGlySerMetAs 1268

Qy       5868 TATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCATGCGGCCAAATGTAAT 5927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 nIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMetArgProAsnValMe 1288

Qy       5928 GGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACAGTCATCGGATGGTAT 5987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 tGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGlnSerSerAspGlyIl 1308

Qy       5988 TGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTGCAACAGCGCTGGACC 6047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 eAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCysAsnSerAlaGlyPr 1328

Qy       6048 ACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACAACACATAAAGAACAT 6107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 oHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGlnHisIleLysAsnIl 1348

Qy       6108 ACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACAAGGGCAGATCCAACT 6167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 eProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGlnGlyGlnIleGlnLe 1368
```

```
Qy       6168 GCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTACTAATAATAATTTAAT 6227
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 uHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProThrAsnAsnAsnLeuMe 1388

Qy       6228 GTCAACTGCCGGAAGTGTCAGTGCTACTAACGGTGTCTCTGGCATCAATTTCGTAGGTCC 6287
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 tSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIleAsnPheValGlyPr 1408

Qy       6288 CTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCAGCAGCAGTTATATGC 6347
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 oSerSerThrAspLeuLysTyrAlaGlnGlnTyrHisSerPheGlnGlnGlnLeuTyrAl 1428

Qy       6348 TACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCACCAGAGCAACATGAT 6407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1428 aThrAsnThrArgSerGlnGlnGlnGlnHisMetHisGlnGlnHisGlnSerAsnMetIl 1448

Qy       6408 AACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTCAACAAA 6456
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 eThrMetProProAsnLeuSerProAsnProThrPhePheValAsnLys 1464
```

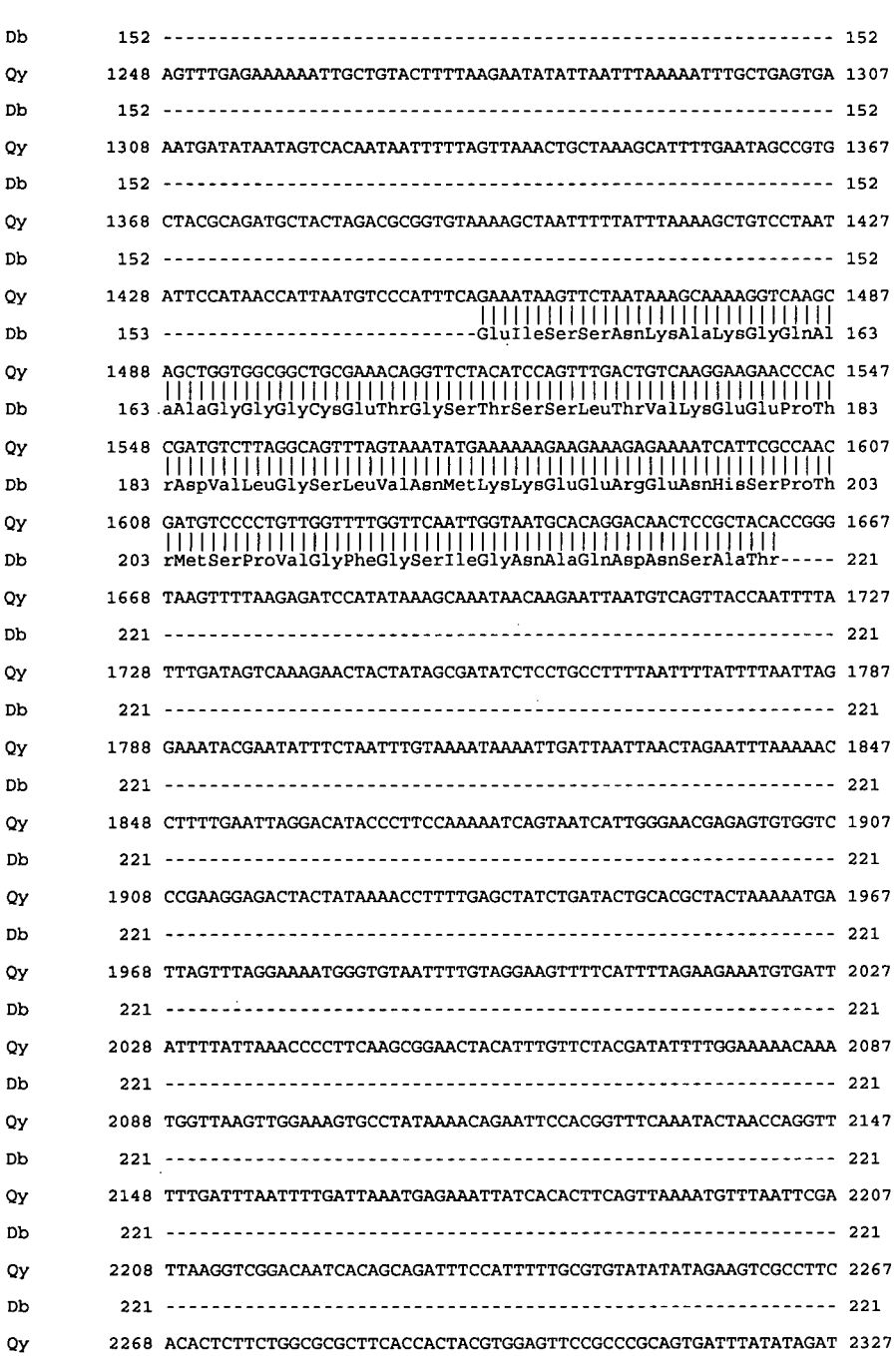
RESULT 3
ADJ71911
ID ADJ71911 standard; protein; 1464 AA.
XX
AC ADJ71911;
XX
DT 20-MAY-2004 (first entry)
XX
DE Fruit fly legless (lgs) polypeptide.
XX
KW Fruit fly; legless; lgs; cell differentiation disorder;
KW cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;
KW breast; head; neck; brain; thyroid; skin; blood disease;
KW tissue regeneration; tissue repair; cytostatic.
XX
OS Drosophila melanogaster.
XX
PN US2004038901-A1.
XX
PD 26-FEB-2004.
XX
PF 22-SEP-2003; 2003US-00664859.
XX
PR 28-JUL-2000; 2000US-0221502P.
PR 27-JUL-2001; 2001US-00915543.
XX
PA (UYZU-) UNIV ZURICH.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2004-203288/19.
DR N-PSDB; ADJ71889.
XX
PT Novel polypeptide sharing one or more homologue amino acid domains with
PT Legless protein being functional homologue of Legless, useful for
PT diagnosing disorders of cell fate.
XX
PS Example 2; SEQ ID NO 1; 62pp; English.
XX
CC The invention relates to a polypeptide sharing one or more homologous
CC amino acid domains with a Legless (Lgs) protein and is therefore a
CC functional homologue of Lgs. The invention also relates to a nucleotide
CC sequence encoding a protein present in invertebrate and/or vertebrate
CC organisms, the nucleotide sequence encoding a protein comprising a
CC positive function in a regulatory pathway and the use of the polypeptide
CC for the isolation of Lgs-binding proteins by carrying out an assay chosen
CC from an in vitro binding assay with such a peptide or a co-
CC immunoprecipitation from vertebrate or invertebrate cell lysates or a
CC mammalian or yeast two hybrid assay. The polypeptide and polynucleotide
CC are useful for treating disorders of cell fate, which involves
CC administering therapeutic compounds chosen from invertebrate and
CC vertebrate Lgs protein homologues or fragments, antibodies, antibody
CC fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA,
CC small peptides or chemical and natural compounds being capable of
CC interfering with Lgs function, synthesis and degradation. The disorders
CC are related to cell differentiation or cell proliferation. The compound
CC is administered to treat a cancerous condition by preventing progression
CC from a pre-neoplastic or non-malignant condition to a neoplastic or
CC malignant state. The cancerous condition is characterised by over-
CC stimulation of the Wnt pathway and is medulloblastoma or cancer of the
CC colon, breast, head and neck, brain, thyroid or skin. The therapeutic
CC compound may also be administered to a blood disease to promote tissue
CC regeneration and repair. This sequence represents the Drosophila Lgs
CC polypeptide of the invention.
XX
SQ Sequence 1464 AA;

Alignment Scores:

Pred. No.:	0	Length:	1464
Score:	7257.00	Matches:	1460
Percent Similarity:	73.01%	Conservative:	1
Best Local Similarity:	72.96%	Mismatches:	3
Query Match:	60.29%	Indels:	537
DB:	8	Gaps:	5

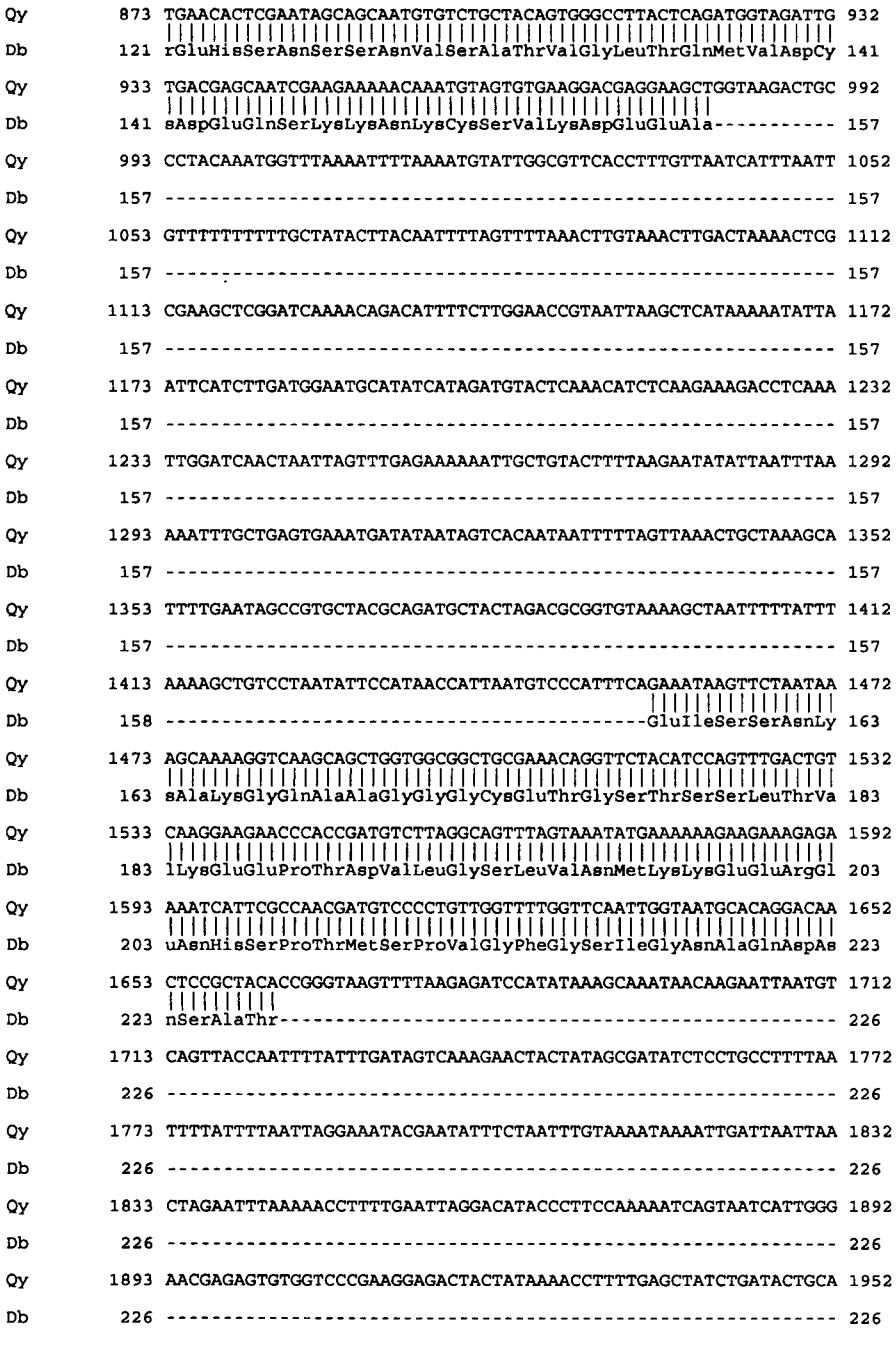
US-10-664-859-1 (1-6909) x ADJ71911 (1-1464)

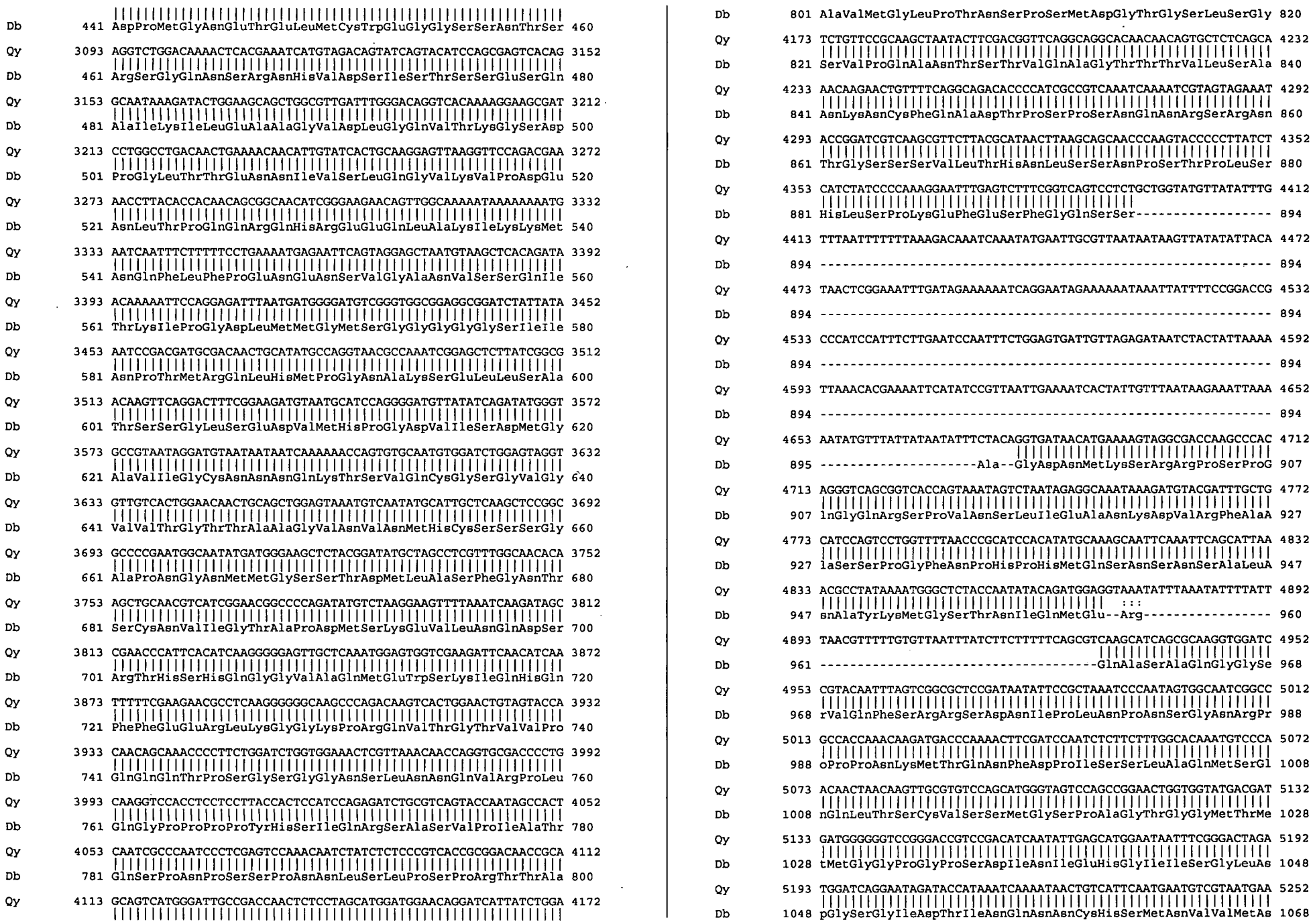
```
Qy        468 ATGCCCCGCAGTCCAACCCAACAACAGCCGCAACCAAACTCCGATGCCTCCTCAACAAGT 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetProArgSerProThrGlnGlnGlnProGlnProAsnSerAspAlaSerSerThrSer 20

Qy        528 GCATCTGGATCAAATCCTGGAGCAGCGATCGGAAATGGGGACTCGGCGGCGAGCAGAAGT 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSerAlaAlaSerArgSer 40

Qy        588 TCTCCGAAGACCCTTAATAGCGAACCCTTTTCTACTTTGTCGCCGGGTAAGACTTGTATT 647
              ||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerProLysThrLeuAsnSerGluProPheSerThrLeu-Ser---------------- 54

Qy        648 GATTTCTCTTTGTCCGGAATTATAACAACTTTCTGTGTTTCCAGATCAAATAAAATTGAC 707
                                                  ||||||||||||||||||||
Db         55 ----------------------------------------ProAspGlnIleLysLeuTh 61

Qy        708 GCCAGAAGAAGGCACTGAGAAAAGCGGACTATCAACTAGTGATAAAGCTGCCACTGGAGG 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 rProGluGluGlyThrGluLysSerGlyLeuSerThrSerAspLysAlaAlaThrGlyGl 81

Qy        768 AGCCCCAGGCAGTGGAAATAATCTGCCCGAGGGACAAACTATGCTAAGGCAGAACTCTAC 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 yAlaProGlySerGlyAsnAsnLeuProGluGlyGlnThrMetLeuArgGlnAsnSerTh 101

Qy        828 GAGCACAATCAACTCGTGCCTAGTCGCTTCTCCACAAAACTCCAGTGAACACTCGAATAG 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 rSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSerGluHisSerAsnSe 121

Qy        888 CAGCAATGTGTCTGCTACAGTGGGCCTTACTCAGATGGTAGATTGTGACGAGCAATCGAA 947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 rSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCysAspGluGlnSerLy 141

Qy        948 GAAAAACAAATGTAGTGTGAAGGACGAGGAAGCTGGTAAGACTGCCCTACAAATGGTTTA 1007
              ||||||||||||||||||||||||||||||||||
Db        141 sLysAsnLysCysSerValLysAspGluGluAla-------------------------- 152

Qy       1008 AAATTTTAAAATGTATTGGCGTTCACCTTTGTTAATCATTTAATTGTTTTTTTTTTGCTA 1067

Db        152 ------------------------------------------------------------ 152

Qy       1068 TACTTACAATTTTAGTTTTAAACTTGTAAACTTGACTAAAACTCGCGAAGCTCGGATCAA 1127

Db        152 ------------------------------------------------------------ 152

Qy       1128 AACAGACATTTTCTTGGAACCGTAATTAAGCTCATAAAAATATTAATTCATCTTGATGGA 1187

Db        152 ------------------------------------------------------------ 152

Qy       1188 ATGCATATCATAGATGTACTCAAACATCTCAAGAAAGACCTCAAATTGGATCAACTAATT 1247
```


Db 152 ---------- 152
Qy 1248 AGTTTGAGAAAAAATTGCTGTACTTTTAAGAATATATTAATTTAAAAATTTGCTGAGTGA 1307
Db 152 ---------- 152
Qy 1308 AATGATATAATAGTCACAATAATTTTTAGTTAAACTGCTAAAGCATTTTGAATAGCCGTG 1367
Db 152 ---------- 152
Qy 1368 CTACGCAGATGCTACTAGACGCGGTGTAAAAGCTAATTTTTATTTAAAAGCTGTCCTAAT 1427
Db 152 ---------- 152
Qy 1428 ATTCCATAACCATTAATGTCCCATTTCAGAAATAAGTTCTAATAAAGCAAAAGGTCAAGC 1487
Db 153 ----------GluIleSerSerAsnLysAlaLysGlyGlnAl 163
Qy 1488 AGCTGGTGGCGGCTGCGAAACAGGTTCTACATCCAGTTTGACTGTCAAGGAAGAACCCAC 1547
Db 163 aAlaGlyGlyGlyCysGluThrGlySerThrSerSerLeuThrValLysGluGluProTh 183
Qy 1548 CGATGTCTTAGGCAGTTTAGTAAATATGAAAAAAGAAGAAAGAGAAAATCATTCGCCAAC 1607
Db 183 rAspValLeuGlySerLeuValAsnMetLysLysGluGluArgGluAsnHisSerProTh 203
Qy 1608 GATGTCCCCTGTTGGTTTTGGTTCAATTGGTAATGCACAGGACAACTCCGCTACACCGGG 1667
Db 203 rMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAsnSerAlaThr----- 221
Qy 1668 TAAGTTTTAAGAGATCCATATAAAGCAAATAACAAGAATTAATGTCAGTTACCAATTTTA 1727
Db 221 ---------- 221
Qy 1728 TTTGATAGTCAAAGAACTACTATAGCGATATCTCCTGCCTTTTAATTTTATTTTAATTAG 1787
Db 221 ---------- 221
Qy 1788 GAAATACGAATATTTCTAATTTGTAAAATAAAATTGATTAATTAACTAGAATTTAAAAAC 1847
Db 221 ---------- 221
Qy 1848 CTTTTGAATTAGGACATACCCTTCCAAAAATCAGTAATCATTGGGAACGAGAGTGTGGTC 1907
Db 221 ---------- 221
Qy 1908 CCGAAGGAGACTACTATAAAACCTTTTGAGCTATCTGATACTGCACGCTACTAAAAATGA 1967
Db 221 ---------- 221
Qy 1968 TTAGTTTAGGAAAATGGGTGTAATTTTGTAGGAAGTTTTCATTTTAGAAGAAATGTGATT 2027
Db 221 ---------- 221
Qy 2028 ATTTTATTAAACCCCTTCAAGCGGAACTACATTTGTTCTACGATATTTTGGAAAAACAAA 2087
Db 221 ---------- 221
Qy 2088 TGGTTAAGTTGGAAAGTGCCTATAAAACAGAATTCCACGGTTTCAAATACTAACCAGGTT 2147
Db 221 ---------- 221
Qy 2148 TTTGATTTAATTTTGATTAAATGAGAAATTATCACACTTCAGTTAAAATGTTTAATTCGA 2207
Db 221 ---------- 221
Qy 2208 TTAAGGTCGGACAATCACAGCAGATTTCCATTTTTGCGTGTATATATAGAAGTCGCCTTC 2267
Db 221 ---------- 221
Qy 2268 ACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCCCGCAGTGATTTATATAGAT 2327
Db 221 ---------- 221
Qy 2328 GATTTACGAGTTATTTAATTTTTTATGGTGTATTTTAATAAATATCTTATTTATTCATTT 2387
Db 221 ---------- 221
Qy 2388 TACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCC 2447
Db 222 -Pro--ValLysIleGluArgIleSerAsnAspSerThrThrGluLysLysGlySerSer 240
Qy 2448 TTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTT 2507
Db 241 LeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGlnLeuAsnProAspPhe 260
Qy 2508 ATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCA 2567
Db 261 IleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuValSerGlyValGlyPro 280
Qy 2568 ATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGA 2627
Db 281 IleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThrAlaAsnAlaAsnGly 300
Qy 2628 ATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCACATATTCGTG 2687
Db 301 IleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGlnAsnHisIlePheVal 320
Qy 2688 TTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTTCAAACT 2747
Db 321 PheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSerGlyGlnPheGlnThr 340
Qy 2748 ATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTTTTTATG 2807
Db 341 IleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeuGluAspPhePheMet 360
Qy 2808 AAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATA 2867
Db 361 LysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerValGlyMetProTrpIle 380
Qy 2868 GGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAG 2927
Db 381 GlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLysIleThrGlnGlnGln 400
Qy 2928 CCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAACAGCAAA 2987
Db 401 ProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGlnHisGluAsnSerLys 420
Qy 2988 CGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGATCCTATGGGCAAC 3047
Db 421 ArgSerThrValSerAlaProSerAsnSerPheValAspGlnSerAspProMetGlyAsn 440
Qy 3048 GAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAAC 3107
Db 441 GluThrGluLeuMetCysTrpGluGlyGlySerSerAsnThrSerArgSerGlyGlnAsn 460
Qy 3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
Db 461 SerArgAsnHisValAspSerIleSerThrSerSerGluSerGlnAlaIleLysIleLeu 480
Qy 3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
Db 481 GluAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAspProGlyLeuThrThr 500
Qy 3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
Db 501 GluAsnAsnIleValSerLeuGlnGlyValLysValProAspGluAsnLeuThrProGln 520
Qy 3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
Db 521 GlnArgGlnHisArgGluGluGlnLeuAlaLysIleLysLysMetAsnGlnPheLeuPhe 540
Qy 3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
Db 541 ProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIleThrLysIleProGly 560

```
Qy      3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       561 AspLeuMetMetGlyMetSerGlyGlyGlyGlySerIleIleAsnProThrMetArg 580

Qy      3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       581 GlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAlaThrSerSerGlyLeu 600

Qy      3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 SerGluAspValMetHisProGlyAspValIleSerAspMetGlyAlaValIleGlyCys 620

Qy      3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       621 AsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGlyValValThrGlyThr 640

Qy      3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       641 ThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGlyAlaProAsnGlyAsn 660

Qy      3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 MetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThrSerCysAsnValIle 680

Qy      3768 GGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACCCATTCACAT 3827
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       681 GlyThrAlaProAspMetSerLysGluValLeuAsnGlnAspSerArgThrHisSerHis 700

Qy      3828 CAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAATTTTTCGAAGAACGC 3887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       701 GlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHisGlnPhePheGluGluArg 720

Qy      3888 CTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACCCCT 3947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 LeuLysGlyGlyLysProArgGlnValThrGlyThrValValProGlnGlnGlnThrPro 740

Qy      3948 TCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCT 4007
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       741 SerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeuGlnGlyProProPro 760

Qy      4008 CCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCC 4067
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       761 ProTyrHisSerIleGlnArgSerAlaSerValProIleAlaThrGlnSerProAsnPro 780

Qy      4068 TCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTG 4127
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 SerSerProAsnAsnLeuSerLeuProSerProArgThrThrAlaAlaValMetGlyLeu 800

Qy      4128 CCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCT 4187
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       801 ProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGlySerValProGlnAla 820

Qy      4188 AATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTT 4247
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       821 AsnThrSerThrValGlnAlaGlyThrThrThrValLeuSerAlaAsnLysAsnCysPhe 840

Qy      4248 CAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGC 4307
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsnThrGlySerSerSer 860

Qy      4308 GTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG 4367
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       861 ValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSerHisLeuSerProLys 880

Qy      4368 GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTTTTTAAA 4427
             ||||||||||||||||||||||||||
Db       881 GluPheGluSerPheGlyGlnSerSer---------------------------------- 889

Qy      4428 GACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGAAATTTG 4487
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       889 ---------------------------------------------------------- 889

Qy      4488 ATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCATTTCTTG 4547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       889 ---------------------------------------------------------- 889

Qy      4548 AATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAATTAAACACGAAAATT 4607
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       889 ---------------------------------------------------------- 889

Qy      4608 CATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATA 4667
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       889 ---------------------------------------------------------- 889

Qy      4668 ATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCAC 4727
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       890 ------Ala--GlyAspAsnMetLysSerArgArgProSerProGlnGlyGlnArgSerP 907

Qy      4728 CAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTT 4787
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       907 roValAsnSerLeuIleGluAlaAsnLysAspValArgPheAlaAlaSerSerProGlyP 927

Qy      4788 TTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGG 4847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       927 heAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuAsnAlaTyrLysMetG 947

Qy      4848 GCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTT 4907
             ||||||||||||||||||||||  :::
Db       947 lySerThrAsnIleGlnMetGlu--Arg------------------------------- 955

Qy      4908 AATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCG 4967
                                     ||||||||||||||||||||||||||||||||||
Db       956 ----------------------GlnAlaSerAlaGlnGlyGlySerValGlnPheSerAr 968

Qy      4968 GCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCCACCAAACAAGAT 5027
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       968 gArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgProProProAsnLysMe 988

Qy      5028 GACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTTG 5087
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       988 tThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGlnGlnLeuThrSerCy 1008

Qy      5088 CGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCGGG 5147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 sValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMetMetGlyGlyProGl 1028

Qy      5148 ACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGATGGATCAGGAATAGA 5207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 yProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAspGlySerGlyIleAs 1048

Qy      5208 TACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAACTCAATGGGTCCCCG 5267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 pThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAsnSerMetGlyProAr 1068

Qy      5268 AATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACCGCCTGGCTTTAATCC 5327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 gMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyProProGlyPheAsnPr 1088

Qy      5328 TAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAA 5387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 oAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerGlyCysGlySerAlaAs 1108

Qy      5388 CTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGT 5447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 nSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMetGlyArgMetProVa 1128

Qy      5448 CAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAGTACCCCAAACACCAT 5507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 lAsnPheGlySerAsnPheAsnProAsnIleGlnValLysAlaSerThrProAsnThrIl 1148

Qy      5508 ACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGGAGCTAATAA 5567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1148 eGlnTyrMetProValArgAlaGlnAsnAlaAsnAsnAsnAsnAsnGlyAlaAsnAs 1168

Qy      5568 TGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGC 5627
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 nValArgMetProProSerLeuGluPheLeuGlnArgTyrAlaAsnProGlnMetGlyAl 1188

Qy       5628 TGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCC 5687
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 aValGlyAsnGlySerProIleCysProProSerAlaSerAspGlyThrProGlyMetPr 1208

Qy       5688 AGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTCTTCCGGAGAGCAACA 5747
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 oGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSerSerGlyGluGlnHi 1228

Qy       5748 CCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTG 5807
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 sGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAsnPhePheGlnAsnCy 1248

Qy       5808 CAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTACCCGGCCATGACGGATCAATGAA 5867
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 sAsnGlnMetSerIleValAspGluGluGlyGlyLeuProGlyHisAspGlySerMetAs 1268

Qy       5868 TATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCATGCGGCCAAATGTAAT 5927
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 nIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMetArgProAsnValMe 1288

Qy       5928 GGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACAGTCATCGGATGGTAT 5987
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 tGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGlnSerSerAspGlyIl 1308

Qy       5988 TGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTGCAACAGCGCTGGACC 6047
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 eAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCysAsnSerAlaGlyPr 1328

Qy       6048 ACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACAACACATAAAGAACAT 6107
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 oHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGlnHisIleLysAsnIl 1348

Qy       6108 ACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACAAGGGCAGATCCAACT 6167
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 eProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGlnGlyGlnIleGlnLe 1368

Qy       6168 GCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTACTAATAATAATTTAAT 6227
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 uHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProThrAsnAsnAsnLeuMe 1388

Qy       6228 GTCAACTGCCGGAAGTGTCAGTGCTACTAACGGTGTCTCTGGCATCAATTTCGTAGGTCC 6287
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 tSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIleAsnPheValGlyPr 1408

Qy       6288 CTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCAGCAGCAGTTATATGC 6347
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 oSerSerThrAspLeuLysTyrAlaGlnGlnTyrHisSerPheGlnGlnGlnLeuTyrAl 1428

Qy       6348 TACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCACCAGAGCAACATGAT 6407
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1428 aThrAsnThrArgSerGlnGlnGlnGlnHisMetHisGlnGlnHisGlnSerAsnMetIl 1448

Qy       6408 AACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTCAACAAA 6456
                    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 eThrMetProProAsnLeuSerProAsnProThrPhePheValAsnLys 1464

RESULT 4
ABB58779
ID   ABB58779 standard; protein; 1429 AA.
XX
AC   ABB58779;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 3129.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02882.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 3129; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1429 AA;

Alignment Scores:
Pred. No.:              0             Length:         1429
Score:                  7040.00       Matches:        1425
Percent Similarity:     71.09%        Conservative:   1
Best Local Similarity:  71.04%        Mismatches:     3
Query Match:            58.49%        Indels:         577
DB:                     4             Gaps:           6

US-10-664-859-1 (1-6909) x ABB58779 (1-1429)

Qy        453 ATGCTCTCGACAACAATGCCCCGCAGTCCAACCCAACAACAGCCGCAACCAAACTCCGAT 512
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetLeuSerThrThrMetProArgSerProThrGlnGlnGlnProGlnProAsnSerAsp 20

Qy        513 GCCTCCTCAACAAGTGCATCTGGATCAAATCCTGGAGCAGCGATCGGAAATGGGGACTCG 572
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaSerSerThrSerAlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSer 40

Qy        573 GCGGCGAGCAGAAGTTCTCCGAAGACCCTTAATAGCGAACCCTTTTCTACTTTGTCGCCG 632
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AlaAlaSerArgSerSerProLysThrLeuAsnSerGluProPheSerThrLeu-Ser-- 59

Qy        633 GGTAAGACTTGTATTGATTTCTCTTTGTCCGGAATTATAACAACTTTCTGTGTTTCCAGA 692
                                                                     |||||
Db         60 ----------------------------------------------------------ProAs 61

Qy        693 TCAAATAAAATTGACGCCAGAAGAAGGCACTGAGAAAAGCGGACTATCAACTAGTGATAA 752
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 pGlnIleLysLeuThrProGluGluGlyThrGluLysSerGlyLeuSerThrSerAspLy 81

Qy        753 AGCTGCCACTGGAGGAGCCCCAGGCAGTGGAAATAATCTGCCCGAGGGACAAACTATGCT 812
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 sAlaAlaThrGlyGlyAlaProGlySerGlyAsnAsnLeuProGluGlyGlnThrMetLe 101

Qy        813 AAGGCAGAACTCTACGAGCACAATCAACTCGTGCCTAGTCGCTTCTCCACAAAACTCCAG 872
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 uArgGlnAsnSerThrSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSe 121
```


Qy 873 TGAACACTCGAATAGCAGCAATGTGTCTGCTACAGTGGGCCTTACTCAGATGGTAGATTG 932
Db 121 rGluHisSerAsnSerSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCy 141
Qy 933 TGACGAGCAATCGAAGAAAAACAAATGTAGTGTGAAGGACGAGGAAGCTGGTAAGACTGC 992
Db 141 sAspGluGlnSerLysLysAsnLysCysSerValLysAspGluGluAla---------- 157
Qy 993 CCTACAAATGGTTTAAAATTTTAAAATGTATTGGCGTTCACCTTTGTTAATCATTTAATT 1052
Db 157 ---------- 157
Qy 1053 GTTTTTTTTTTGCTATACTTACAATTTTAGTTTTAAACTTGTAAACTTGACTAAAACTCG 1112
Db 157 ---------- 157
Qy 1113 CGAAGCTCGGATCAAAACAGACATTTTCTTGGAACCGTAATTAAGCTCATAAAAATATTA 1172
Db 157 ---------- 157
Qy 1173 ATTCATCTTGATGGAATGCATATCATAGATGTACTCAAACATCTCAAGAAAGACCTCAAA 1232
Db 157 ---------- 157
Qy 1233 TTGGATCAACTAATTAGTTTGAGAAAAAATTGCTGTACTTTTAAGAATATATTAATTTAA 1292
Db 157 ---------- 157
Qy 1293 AAATTTGCTGAGTGAAATGATATAATAGTCACAATAATTTTTAGTTAAACTGCTAAAGCA 1352
Db 157 ---------- 157
Qy 1353 TTTTGAATAGCCGTGCTACGCAGATGCTACTAGACGCGGTGTAAAAGCTAATTTTTATTT 1412
Db 157 ---------- 157
Qy 1413 AAAAGCTGTCCTAATATTCCATAACCATTAATGTCCCATTTCAGAAATAAGTTCTAATAA 1472
Db 158 ----------GluIleSerSerAsnLy 163
Qy 1473 AGCAAAAGGTCAAGCAGCTGGTGGCGGCTGCGAAACAGGTTCTACATCCAGTTTGACTGT 1532
Db 163 sAlaLysGlyGlnAlaAlaGlyGlyGlyCysGluThrGlySerThrSerSerLeuThrVa 183
Qy 1533 CAAGGAAGAACCCACCGATGTCTTAGGCAGTTTAGTAAATATGAAAAAAGAAGAAAGAGA 1592
Db 183 lLysGluGluProThrAspValLeuGlySerLeuValAsnMetLysLysGluGluArgGl 203
Qy 1593 AAATCATTCGCCAACGATGTCCCCTGTTGGTTTTGGTTCAATTGGTAATGCACAGGACAA 1652
Db 203 uAsnHisSerProThrMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAs 223
Qy 1653 CTCCGCTACACCGGGTAAGTTTTAAGAGATCCATATAAAGCAAATAACAAGAATTAATGT 1712
Db 223 nSerAlaThr---------- 226
Qy 1713 CAGTTACCAATTTTATTTGATAGTCAAAGAACTACTATAGCGATATCTCCTGCCTTTTAA 1772
Db 226 ---------- 226
Qy 1773 TTTTATTTTAATTAGGAAATACGAATATTTCTAATTTGTAAAATAAAATTGATTAATTAA 1832
Db 226 ---------- 226
Qy 1833 CTAGAATTTAAAAACCTTTTGAATTAGGACATACCCTTCCAAAAATCAGTAATCATTGGG 1892
Db 226 ---------- 226
Qy 1893 AACGAGAGTGTGGTCCCGAAGGAGACTACTATAAAACCTTTTGAGCTATCTGATACTGCA 1952
Db 226 ---------- 226
Qy 1953 CGCTACTAAAAATGATTAGTTTAGGAAAATGGGTGTAATTTTGTAGGAAGTTTTCATTTT 2012
Db 226 ---------- 226
Qy 2013 AGAAGAAATGTGATTATTTTATTAAACCCCTTCAAGCGGAACTACATTTGTTCTACGATA 2072
Db 226 ---------- 226
Qy 2073 TTTTGGAAAAACAAATGGTTAAGTTGGAAAGTGCCTATAAAACAGAATTCCACGGTTTCA 2132
Db 226 ---------- 226
Qy 2133 AATACTAACCAGGTTTTTGATTTAATTTTGATTAAATGAGAAATTATCACACTTCAGTTA 2192
Db 226 ---------- 226
Qy 2193 AAATGTTTAATTCGATTAAGGTCGGACAATCACAGCAGATTTCCATTTTTGCGTGTATAT 2252
Db 226 ---------- 226
Qy 2253 ATAGAAGTCGCCTTCACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCCCGCA 2312
Db 226 ---------- 226
Qy 2313 GTGATTTATATAGATGATTTACGAGTTATTTAATTTTTTATGGTGTATTTTAATAAATAT 2372
Db 226 ---------- 226
Qy 2373 CTTATTTATTCATTTTACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAA 2432
Db 227 ----------Pro--ValLysIleGluArgIleSerAsnAspSerThrThrGlu 240
Qy 2433 AAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAG 2492
Db 241 LysLysGlySerSerLeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGln 260
Qy 2493 TTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTA 2552
Db 261 LeuAsnProAspPheIleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuVal 280
Qy 2553 AGCGGAGTAGGACCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACT 2612
Db 281 SerGlyValGlyProIleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThr 300
Qy 2613 GCCAACGCCAATGGAATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAA 2672
Db 301 AlaAsnAlaAsnGlyIleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGln 320
Qy 2673 AATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGC 2732
Db 321 AsnHisIlePheValPheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSer 340
Qy 2733 GGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTG 2792
Db 341 GlyGlnPheGlnThrIleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeu 360
Qy 2793 GAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTC 2852
Db 361 GluAspPhePheMetLysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerVal 380
Qy 2853 GGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAA 2912
Db 381 GlyMetProTrpIleGlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLys 400
Qy 2913 ATAACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAA 2972
Db 401 IleThrGlnGlnGlnProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGln 420
Qy 2973 CATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCT 3032
Db 421 HisGluAsnSerLysArgSerThrValSerAlaProSerAsnSerPheValAspGlnSer 440
Qy 3033 GATCCTATGGGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT 3092


Db 441 AspProMetGlyAsnGluThrGluLeuMetCysTrpGluGlyGlySerSerAsnThrSer 460
Qy 3093 AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAG 3152
Db 461 ArgSerGlyGlnAsnSerArgAsnHisValAspSerIleSerThrSerSerGluSerGln 480
Qy 3153 GCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGAT 3212
Db 481 AlaIleLysIleLeuGluAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAsp 500
Qy 3213 CCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAA 3272
Db 501 ProGlyLeuThrThrGluAsnAsnIleValSerLeuGlnGlyValLysValProAspGlu 520
Qy 3273 AACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATG 3332
Db 521 AsnLeuThrProGlnGlnArgGlnHisArgGluGluGlnLeuAlaLysIleLysLysMet 540
Qy 3333 AATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATA 3392
Db 541 AsnGlnPheLeuPheProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIle 560
Qy 3393 ACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATA 3452
Db 561 ThrLysIleProGlyAspLeuMetMetGlyMetSerGlyGlyGlyGlySerIleIle 580
Qy 3453 AATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCG 3512
Db 581 AsnProThrMetArgGlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAla 600
Qy 3513 ACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGT 3572
Db 601 ThrSerSerGlyLeuSerGluAspValMetHisProGlyAspValIleSerAspMetGly 620
Qy 3573 GCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGT 3632
Db 621 AlaValIleGlyCysAsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGly 640
Qy 3633 GTTGTCACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGC 3692
Db 641 ValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGly 660
Qy 3693 GCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACA 3752
Db 661 AlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThr 680
Qy 3753 AGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGC 3812
Db 681 SerCysAsnValIleGlyThrAlaProAspMetSerLysGluValLeuAsnGlnAspSer 700
Qy 3813 CGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
Db 701 ArgThrHisSerHisGlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHisGln 720
Qy 3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
Db 721 PhePheGluGluArgLeuLysGlyGlyLysProArgGlnValThrGlyThrValValPro 740
Qy 3933 CAACAGCAAACCCCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTG 3992
Db 741 GlnGlnGlnThrProSerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeu 760
Qy 3993 CAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACT 4052
Db 761 GlnGlyProProProProTyrHisSerIleGlnArgSerAlaSerValProIleAlaThr 780
Qy 4053 CAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCA 4112
Db 781 GlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThrAla 800
Qy 4113 GCAGTCATGGGATTGCCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGA 4172
Db 801 AlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGly 820
Qy 4173 TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCA 4232
Db 821 SerValProGlnAlaAsnThrSerThrValGlnAlaGlyThrThrThrValLeuSerAla 840
Qy 4233 AACAAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAAT 4292
Db 841 AsnLysAsnCysPheGlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsn 860
Qy 4293 ACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCT 4352
Db 861 ThrGlySerSerSerValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSer 880
Qy 4353 CATCTATCCCCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTG 4412
Db 881 HisLeuSerProLysGluPheGluSerPheGlyGlnSerSer---------- 894
Qy 4413 TTTAATTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACA 4472
Db 894 ---------- 894
Qy 4473 TAACTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG 4532
Db 894 ---------- 894
Qy 4533 CCCATCCATTTCTTGAATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAA 4592
Db 894 ---------- 894
Qy 4593 TTAAACACGAAAATTCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAA 4652
Db 894 ---------- 894
Qy 4653 AATATGTTTATTATAATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCAC 4712
Db 895 ----------Ala--GlyAspAsnMetLysSerArgArgProSerProG 907
Qy 4713 AGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTG 4772
Db 907 lnGlyGlnArgSerProValAsnSerLeuIleGluAlaAsnLysAspValArgPheAlaA 927
Qy 4773 CATCCAGTCCTGGTTTTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAA 4832
Db 927 laSerSerProGlyPheAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuA 947
Qy 4833 ACGCCTATAAAATGGGCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATT 4892
Db 947 snAlaTyrLysMetGlySerThrAsnIleGlnMetGlu--Arg---------- 960
Qy 4893 TAACGTTTTTGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATC 4952
Db 961 ----------GlnAlaSerAlaGlnGlyGlySe 968
Qy 4953 CGTACAATTTAGTCGGCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCC 5012
Db 968 rValGlnPheSerArgArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgPr 988
Qy 5013 GCCACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCA 5072
Db 988 oProProAsnLysMetThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGl 1008
Qy 5073 ACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGAT 5132
Db 1008 nGlnLeuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMe 1028
Qy 5133 GATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGA 5192
Db 1028 tMetGlyGlyProGlyProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAs 1048
Qy 5193 TGGATCAGGAATAGATACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAA 5252
Db 1048 pGlySerGlyIleAspThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAs 1068

```
Qy       5253 CTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACC 5312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 nSerMetGlyProArgMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyPr 1088

Qy       5313 GCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGG 5372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1088 oProGlyPheAsnProAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerGl 1108

Qy       5373 CTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGAT 5432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 yCysGlySerAlaAsnSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMe 1128

Qy       5433 GGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAG 5492
             ||||||||||||||||||||||||||||||||||||||||||||
Db       1128 tGlyArgMetProValAsnPheGlySerAsnPheAsnProAsnIle---------------- 1143

Qy       5493 TACCCCAAACACCATACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAA 5552

Db       1143 ------------------------------------------------------------ 1143

Qy       5553 CAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAA 5612
                                                          ||||||||||||||
Db       1144 ----------------------------------------------GlnArgTyrAlaAs 1148

Qy       5613 CCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGG 5672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1148 nProGlnMetGlyAlaValGlyAsnGlySerProIleCysProProSerAlaSerAspGl 1168

Qy       5673 TACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTC 5732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 yThrProGlyMetProGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSe 1188

Qy       5733 TTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAA 5792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 rSerGlyGluGlnHisGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAs 1208

Qy       5793 CTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTACCCGGCCA 5852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 nPhePheGlnAsnCysAsnGlnMetSerIleValAspGluGluGlyGlyLeuProGlyHi 1228

Qy       5853 TGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCAT 5912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 sAspGlySerMetAsnIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMe 1248

Qy       5913 GCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACA 5972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 tArgProAsnValMetGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGl 1268

Qy       5973 GTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTG 6032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 nSerSerAspGlyIleAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCy 1288

Qy       6033 CAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACA 6092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 sAsnSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGl 1308

Qy       6093 ACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACA 6152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 nHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGl 1328

Qy       6153 AGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTAC 6212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 nGlyGlnIleGlnLeuHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProTh 1348

Qy       6213 TAATAATAATTTAATGTCAACTGCCGGAAGTGTCAGTGCTACTAACGGTGTCTCTGGCAT 6272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 rAsnAsnAsnLeuMetSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIl 1368

Qy       6273 CAATTTCGTAGGTCCCTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCA 6332
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 eAsnPheValGlyProSerSerThrAspLeuLysTyrAlaGlnGlnTyrHisSerPheGl 1388

Qy       6333 GCAGCAGTTATATGCTACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCA 6392
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 nGlnGlnLeuTyrAlaThrAsnThrArgSerGlnGlnGlnGlnHisMetHisGlnGlnHi 1408

Qy       6393 CCAGAGCAACATGATAACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTCAA 6452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 sGlnSerAsnMetIleThrMetProProAsnLeuSerProAsnProThrPhePheValAs 1428

Qy       6453 CAAA 6456
             ||||
Db       1428 nLys 1429
```

RESULT 5
ABW01532
ID ABW01532 standard; peptide; 112 AA.
XX
AC ABW01532;
XX
DT 15-JAN-2004 (first entry)
XX
DE Drosophila species legless (lgs) peptide #5.
XX
KW Legless protein; lgs; cell fate disorder; blood disease; gene therapy;
KW cancer; tissue regeneration; tissue repair; cytostatic.
XX
OS Drosophila sp.
XX
PN US2003114413-A1.
XX
PD 19-JUN-2003.
XX
PF 19-DEC-2002; 2002US-00322579.
XX
PR 28-JUL-2000; 2000US-0221502P.
PR 27-JUL-2001; 2001US-00915543.
XX
PA (UYZU-) UNIV ZURICH.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2003-829432/77.
XX
PT Novel Lgs polypeptide useful for isolation of Lgs-binding proteins,
PT diagnosing disorders of cell fate, treating diseases such as cancer.
XX
PS Claim 28; Fig 7B; 0pp; English.
XX
CC The invention relates to novel legless (lgs) proteins and polynucleotides
CC encoding such proteins. Lgs sequences are useful for the treatment of
CC disorders of cell fate such as differentiation or proliferation. The
CC invention is used to treat blood disease or a cancerous condition
CC characterised by over-stimulation of the Wnt pathway such as colon,
CC breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and
CC is administered to prevent progression from a pre-neoplastic or non-
CC malignant condition to a neoplastic or malignant state. It is
CC administered to promote tissue regeneration and repair. The invention is
CC also useful in the therapy of diseases cost by an over-activation of Wg
CC pathway. It is useful for reducing lgs gene expression in an invertebrate
CC or vertebrate organism or an invertebrate or vertebrate cell line. The
CC invention is also useful in gene therapy. The present sequence is
CC Drosophila species legless (lgs) peptide
XX
SQ Sequence 112 AA;

Alignment Scores:

Pred. No.:	9.01e-40	Length:	112
Score:	570.00	Matches:	112
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.74%	Indels:	0
DB:	7	Gaps:	0

US-10-664-859-1 (1-6909) x ABW01532 (1-112)

```
Qy       4029 TCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCT 4088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnAsnLeuSer 20

Qy       4089 CTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATG 4148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40

Qy       4149 GATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCA 4208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60

Qy       4209 GGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCG 4268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlyThrThrThrValLeuSerAlaAsnLysAsnCysPheGlnAlaAspThrProSerPro 80

Qy       4269 TCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGC 4328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerAsnGlnAsnArgSerArgAsnThrGlySerSerSerValLeuThrHisAsnLeuSer 100

Qy       4329 AGCAACCCAAGTACCCCCTTATCTCATCTATCCCCA 4364
             |||||||||||||||||||||||||||||||||||||
Db        101 SerAsnProSerThrProLeuSerHisLeuSerPro 112
```

RESULT 6
ADJ71898
ID ADJ71898 standard; protein; 112 AA.
XX
AC ADJ71898;
XX
DT 20-MAY-2004 (first entry)
XX
DE Fruit fly legless (lgs) peptide fragment #5.
XX
KW Fruit fly; legless; lgs; cell differentiation disorder;
KW cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;
KW breast; head; neck; brain; thyroid; skin; blood disease;
KW tissue regeneration; tissue repair; cytostatic.
XX
OS Drosophila melanogaster.
XX
PN US2004038901-A1.
XX
PD 26-FEB-2004.
XX
PF 22-SEP-2003; 2003US-00664859.
XX
PR 28-JUL-2000; 2000US-0221502P.
PR 27-JUL-2001; 2001US-00915543.
XX
PA (UYZU-) UNIV ZURICH.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2004-203288/19.
XX
PT Novel polypeptide sharing one or more homologue amino acid domains with
PT Legless protein being functional homologue of Legless, useful for
PT diagnosing disorders of cell fate.
XX
PS Disclosure; SEQ ID NO 10; 62pp; English.
XX
CC The invention relates to a polypeptide sharing one or more homologous
CC amino acid domains with a Legless (Lgs) protein and is therefore a
CC functional homologue of Lgs. The invention also relates to a nucleotide
CC sequence encoding a protein present in invertebrate and/or vertebrate
CC organisms, the nucleotide sequence encoding a protein comprising a
CC positive function in a regulatory pathway and the use of the polypeptide
CC for the isolation of Lgs-binding proteins by carrying out an assay chosen
CC from an in vitro binding assay with such a peptide or a co-
CC immunoprecipitation from vertebrate or invertebrate cell lysates or a
CC mammalian or yeast two hybrid assay. The polypeptide and polynucleotide
CC are useful for treating disorders of cell fate, which involves
CC administering therapeutic compounds chosen from invertebrate and
CC vertebrate Lgs protein homologues or fragments, antibodies, antibody
CC fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA,
CC small peptides or chemical and natural compounds being capable of
CC interfering with Lgs function, synthesis and degradation. The disorders
CC are related to cell differentiation or cell proliferation. The compound
CC is administered to treat a cancerous condition by preventing progression
CC from a pre-neoplastic or non-malignant condition to a neoplastic or
CC malignant state. The cancerous condition is characterised by over-
CC stimulation of the Wnt pathway and is medulloblastoma or cancer of the
CC colon, breast, head and neck, brain, thyroid or skin. The therapeutic
CC compound may also be administered to a blood disease to promote tissue
CC regeneration and repair. This sequence represents a Drosophila Lgs
CC peptide fragment of the invention.
XX
SQ Sequence 112 AA;

Alignment Scores:

Pred. No.:	9.01e-40	Length:	112
Score:	570.00	Matches:	112
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.74%	Indels:	0
DB:	8	Gaps:	0

US-10-664-859-1 (1-6909) x ADJ71898 (1-112)

```
Qy       4029 TCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCT 4088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SerAlaSerValProIleAlaThrGlnSerProAsnProSerSerProAsnAsnLeuSer 20

Qy       4089 CTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATG 4148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 LeuProSerProArgThrThrAlaAlaValMetGlyLeuProThrAsnSerProSerMet 40

Qy       4149 GATGGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCA 4208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AspGlyThrGlySerLeuSerGlySerValProGlnAlaAsnThrSerThrValGlnAla 60

Qy       4209 GGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCG 4268
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GlyThrThrThrValLeuSerAlaAsnLysAsnCysPheGlnAlaAspThrProSerPro 80

Qy       4269 TCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGC 4328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerAsnGlnAsnArgSerArgAsnThrGlySerSerSerValLeuThrHisAsnLeuSer 100

Qy       4329 AGCAACCCAAGTACCCCCTTATCTCATCTATCCCCA 4364
             |||||||||||||||||||||||||||||||||||||
Db        101 SerAsnProSerThrProLeuSerHisLeuSerPro 112
```

RESULT 7
ABB11808
ID ABB11808 standard; peptide; 1435 AA.
XX
AC ABB11808;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human BCL9 homologue, SEQ ID NO:2178.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;

KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA09052.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 20; Page 256-257; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention
XX
SQ Sequence 1435 AA;

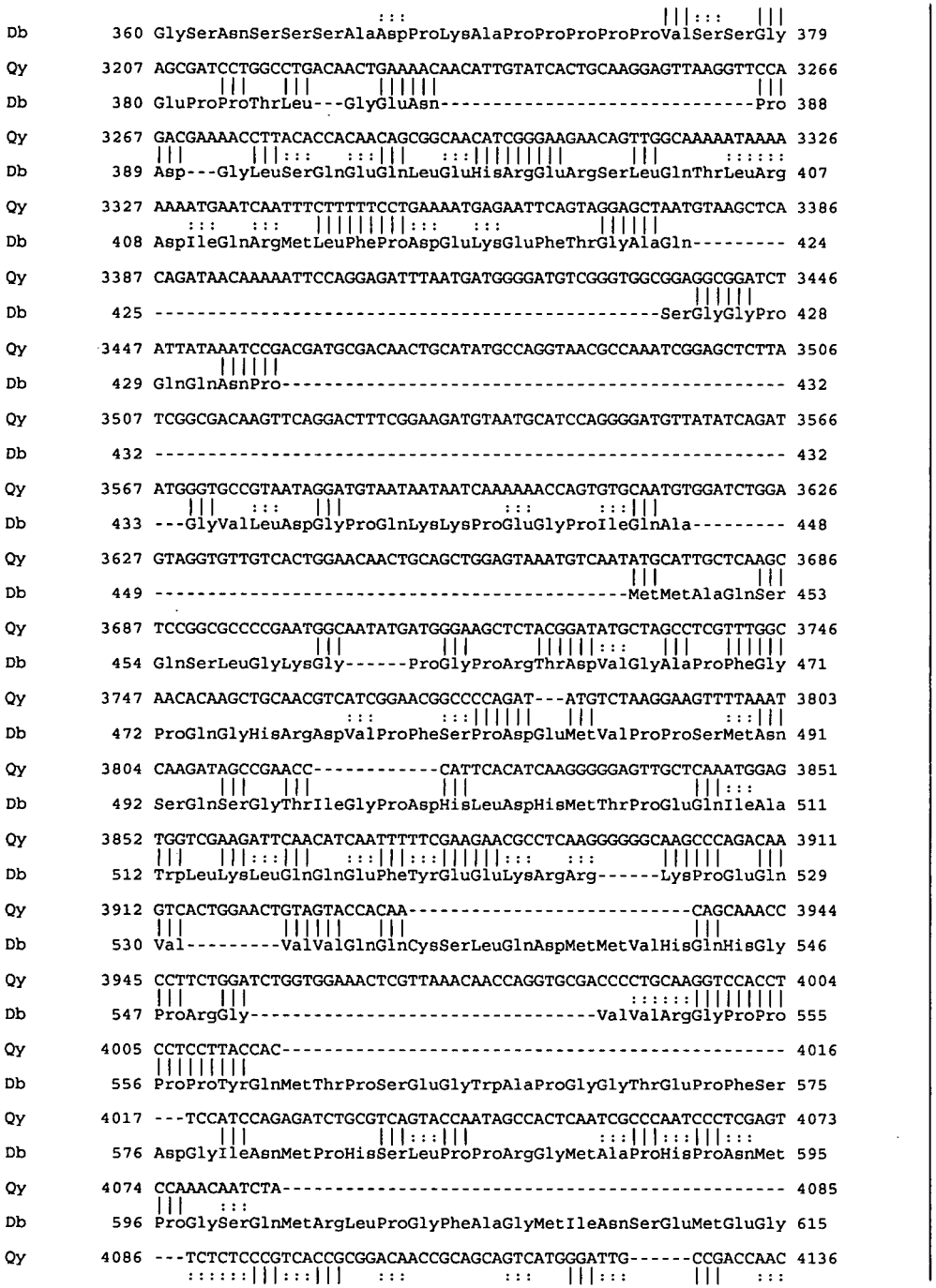
Alignment Scores:

Pred. No.:	7.17e-18	Length:	1435
Score:	321.50	Matches:	313
Percent Similarity:	30.10%	Conservative:	180
Best Local Similarity:	19.11%	Mismatches:	522
Query Match:	2.67%	Indels:	624

DB: 4 Gaps: 71

US-10-664-859-1 (1-6909) x ABB11808 (1-1435)

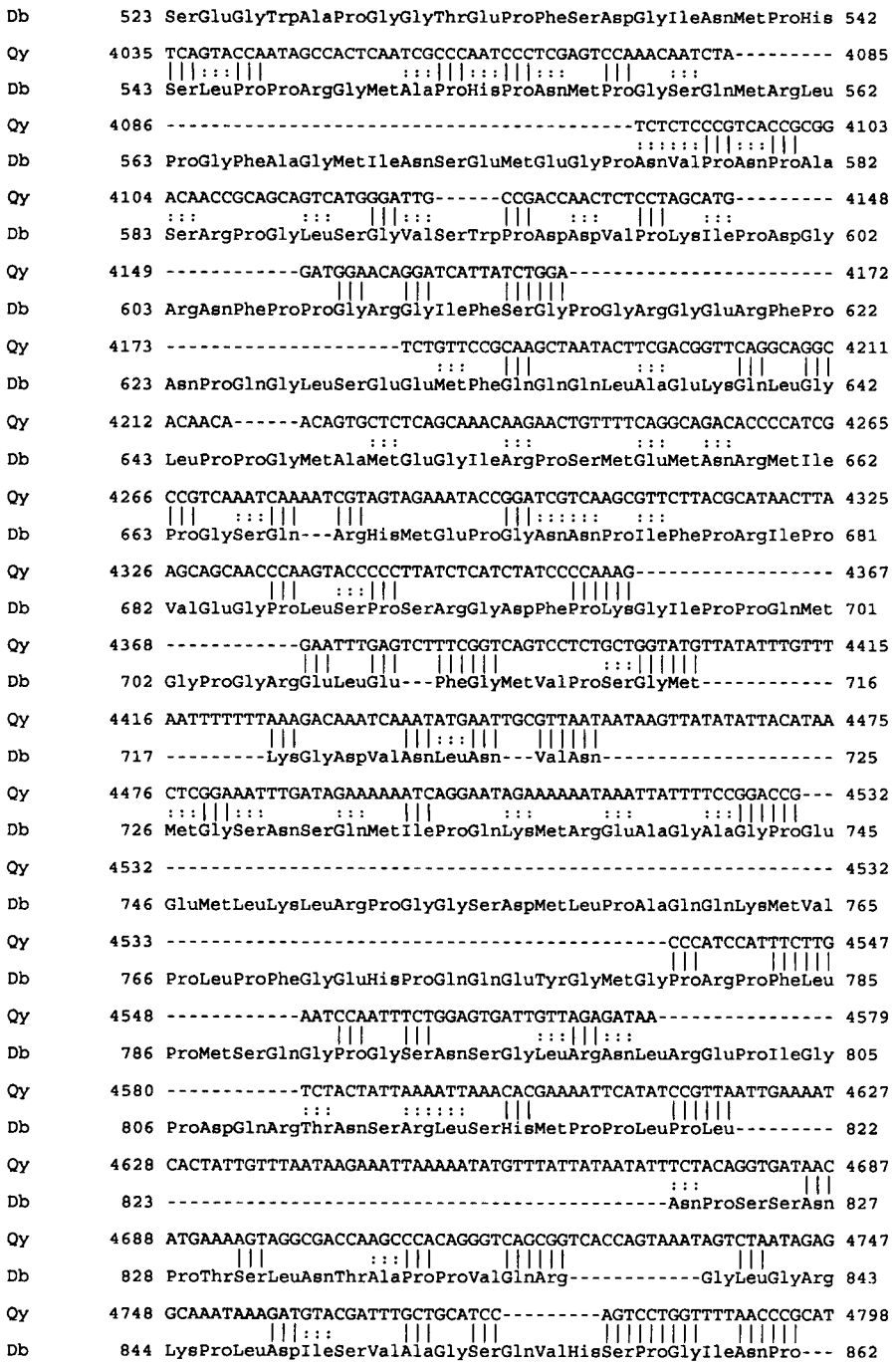
```
Qy      2412 TCAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATG 2471
             :::      |||:::   :::   |||         :::                  |||
Db        55 AsnThrGlnSerSerProLysSerLysGlnGluValMetValArgProProThrValMet  74

Qy      2472 AGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCT 2531
             |||      |||      ||||||:::      |||   |||:::
Db        75 SerProSerGlyAsnProGlnLeuAspSerLysPheSerAsnGlnGlyLysGlnGlyGly  94

Qy      2532 GCAATTTCGAGCATATTAGTAAGC---------------------------------- 2555
             :::   |||      |||
Db        95 SerAlaSerGlnSerGlnProSerProCysAspSerLysSerGlyGlyHisThrProLys 114

Qy      2556 ---------GGAGTAGGACCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTA 2606
                     |||   |||         |||:::      |||||||||   |||
Db       115 AlaLeuProGlyProGlyGlySerMetGlyLeuLysAsnGlyAlaGlyAsnGly------ 132

Qy      2607 TTGACTGCCAACGCCAATGGA--------------ATCTCCTCGGGTAGC--------- 2642
                    |||          |||              ||||||:::   |||
Db       133 ------AlaLysGlyLysGlyLysArgGluArgSerIleSerAlaAspSerPheAspGln 150

Qy      2643 ----------------------------------------AGTAATTGTTTGGATTAC 2660
                                                     |||     |||:::
Db       151 ArgAspProGlyThrProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHis 170

Qy      2661 ATGCAACAGCAAAATCAC---------------------------------------- 2678
             ::::::   |||:::
Db       171 IleLysSerGlnAspSerGlnHisThrProHisSerMetThrProSerAsnAlaThrAla 190

Qy      2678 ---------------------------------------------------------- 2678

Db       191 ProArgSerSerThrProProHisGlyGlnThrThrAlaThrGluProThrProAlaGln 210

Qy      2679 ------------------ATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAA 2720
                               ::::::||||||||||||:::::|||||||||   ||||||
Db       211 LysThrProAlaLysValValTyrValPheSerThrGluMetAlaAsnLysAlaAlaGlu 230

Qy      2721 TCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACA 2780
             :::||||||   ||||||   :::||||||:::::::::|||
Db       231 AlaValLeuLysGlyGlnValGluThrIleValSerPheHis------------------ 244

Qy      2781 AAAAGCTTCCTGGAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGG 2840
                                     ::::::|||   :::      ||||||   :::|||
Db       245 ----------------------IleGlnAsn---IleSerAsnAsnLysThrGluArg 255

Qy      2841 CACAATTCCGTCGGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAAT 2900
                    :::      |||                                      |||
Db       256 SerThrAla---------ProLeuAsnThrGlnIleSerAlaLeuArgAsnAspProLys 272

Qy      2901 CCTGTAGCCAAAATAACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCC 2960
             |||:::               |||||||||
Db       273 ProLeu------------ProGlnGlnProProAlaProAla----------------- 282

Qy      2961 CAATTCAATCAACATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT 3020
                    ||||||   :::||||||   ::::::|||         |||:::
Db       283 ------AsnGlnAspGlnAsnSerSerGlnAsnThrArgLeuGlnProThrProProIle 300

Qy      3021 ------------------------------------GTCGACCAGTCTGATCCTATG 3041
                                                 :::|||:::      |||
Db       301 ProAlaProAlaProLysProAlaAlaProProArgProLeuAspArgGluSerPro--- 319

Qy      3042 GGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT--------- 3092
             |||   |||   :::|||:::         |||      ::::::::::::
Db       320 GlyValGluAsnLysLeuIleProSerValGlySerProAlaSerSerThrProLeuPro 339

Qy      3093 ------AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAG 3146
                   :::|||   ||||||      :::      :::::::::      |||:::
Db       340 ProAspGlyThrGlyProAsnSerThrProAsnAsnArgAlaValThrProValSerGln 359

Qy      3147 TCACAGGCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGA 3206
```


Db 360 GlySerAsnSerSerSerAlaAspProLysAlaProProProProProValSerSerGly 379
Qy 3207 AGCGATCCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCA 3266
Db 380 GluProProThrLeu---GlyGluAsn----------Pro 388
Qy 3267 GACGAAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAA 3326
Db 389 Asp---GlyLeuSerGlnGluGlnLeuGluHisArgGluArgSerLeuGlnThrLeuArg 407
Qy 3327 AAAATGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCA 3386
Db 408 AspIleGlnArgMetLeuPheProAspGluLysGluPheThrGlyAlaGln--------- 424
Qy 3387 CAGATAACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCT 3446
Db 425 ----------SerGlyGlyPro 428
Qy 3447 ATTATAAATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTA 3506
Db 429 GlnGlnAsnPro---------- 432
Qy 3507 TCGGCGACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGAT 3566
Db 432 ---------- 432
Qy 3567 ATGGGTGCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGA 3626
Db 433 ---GlyValLeuAspGlyProGlnLysLysProGluGlyProIleGlnAla--------- 448
Qy 3627 GTAGGTGTTGTCACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGC 3686
Db 449 ----------MetMetAlaGlnSer 453
Qy 3687 TCCGGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGC 3746
Db 454 GlnSerLeuGlyLysGly------ProGlyProArgThrAspValGlyAlaProPheGly 471
Qy 3747 AACACAAGCTGCAACGTCATCGGAACGGCCCCAGAT---ATGTCTAAGGAAGTTTTAAAT 3803
Db 472 ProGlnGlyHisArgAspValProPheSerProAspGluMetValProProSerMetAsn 491
Qy 3804 CAAGATAGCCGAACC----------CATTCACATCAAGGGGGAGTTGCTCAAATGGAG 3851
Db 492 SerGlnSerGlyThrIleGlyProAspHisLeuAspHisMetThrProGluGlnIleAla 511
Qy 3852 TGGTCGAAGATTCAACATCAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAA 3911
Db 512 TrpLeuLysLeuGlnGlnGluPheTyrGluGluLysArgArg------LysProGluGln 529
Qy 3912 GTCACTGGAACTGTAGTACCACAA----------CAGCAAACC 3944
Db 530 Val---------ValValGlnGlnCysSerLeuGlnAspMetMetValHisGlnHisGly 546
Qy 3945 CCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCT 4004
Db 547 ProArgGly----------ValValArgGlyProPro 555
Qy 4005 CCTCCTTACCAC---------- 4016
Db 556 ProProTyrGlnMetThrProSerGluGlyTrpAlaProGlyGlyThrGluProPheSer 575
Qy 4017 ---TCCATCCAGAGATCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGT 4073
Db 576 AspGlyIleAsnMetProHisSerLeuProProArgGlyMetAlaProHisProAsnMet 595
Qy 4074 CCAAACAATCTA---------- 4085
Db 596 ProGlySerGlnMetArgLeuProGlyPheAlaGlyMetIleAsnSerGluMetGluGly 615
Qy 4086 ---TCTCTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTG------CCGACCAAC 4136
Db 616 ProAsnValProAsnProAlaSerArgProGlyLeuSerGlyValSerTrpProAspAsp 635
Qy 4137 TCTCCTAGCATG----------GATGGAACAGGATCATTATCTGGA--- 4172
Db 636 ValProLysIleProAspGlyArgAsnPheProProGlyGlnGlyIlePheSerGlyPro 655
Qy 4173 ----------TCTGTTCCGCAAGCTAAT 4190
Db 656 GlyArgGlyGluArgPheProAsnProGlnGlyLeuSerGluGluMetPheGlnGlnGln 675
Qy 4191 ACTTCGACGGTTCAGGCAGGCACAACA------ACAGTGCTCTCAGCAAACAAGAACTGT 4244
Db 676 LeuAlaGluLysGlnLeuGlyLeuProProGlyMetAlaMetGluGlyIleArgProSer 695
Qy 4245 TTTCAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCA 4304
Db 696 MetGluMetAsnArgMetIleProGlySerGln---ArgHisMetGluProGlyAsnAsn 714
Qy 4305 AGCGTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCA 4364
Db 715 ProIleProProArgIleProValGluGlyProLeuSerProSerArgGlyAspPhePro 734
Qy 4365 AAG----------GAATTTGAGTCTTTCGGTCAGTCCTCT 4394
Db 735 LysGlyIleProProGlnMetGlyProGlyArgGluLeuGlu---PheGlyMetValPro 753
Qy 4395 GCTGGTATGTTATATTTGTTTAATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAAT 4454
Db 754 SerGlyMet----------LysGlyAspValAsnLeuAsn---ValAsn 765
Qy 4455 AATAAGTTATATATTACATAACTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAAT 4514
Db 766 ----------MetGlySerAsnSerGlnMetIleProGlnLysMetArg 778
Qy 4515 AAATTATTTTCCGGACCG---------- 4532
Db 779 GluAlaGlyAlaGlyProGluGluMetLeuLysLeuArgProGlyGlySerAspMetLeu 798
Qy 4532 ---------- 4532
Db 799 ProAlaGlnGlnLysMetValProLeuProPheGlyGluHisProGlnGlnGluTyrGly 818
Qy 4533 ------CCCATCCATTTCTTG----------AATCCAATTTCTGGAGTGATTGTTAGA 4574
Db 819 MetGlyProArgProPheLeuProMetSerGlnGlyProGlySerAsnSerGlyLeuArg 838
Qy 4575 GATAA----------TCTACTATTAAAATTAAACACGAAAAT 4606
Db 839 AsnLeuArgGluProIleGlyProAspGlnArgThrAsnSerArgLeuSerHisMetPro 858
Qy 4607 TCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTTATTAT 4666
Db 859 ProLeuProLeu---------- 862
Qy 4667 AATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCA 4726
Db 863 ------AsnProSerSerAsnProThrSerLeuAsnThrAlaProProValGlnArg--- 879
Qy 4727 CCAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCC--------- 4777
Db 880 ---------GlyLeuGlyArgLysProLeuAspIleSerValAlaGlySerGlnValHis 896
Qy 4778 AGTCCTGGTTTTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCC 4837
Db 897 SerProGlyIleAsnPro---------LeuLysSerProThrMetHisGlnValGlnSer 913
Qy 4838 TATAAAATGGGCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATTTAACG 4897
Db 914 ProMetLeuGlySerProSerGlyAsnLeu---------- 923
Qy 4898 TTTTTGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTAC 4957
Db 924 ----------LysSerProGlnThrProSer 930

```
Qy       4958 AATTTAGTCGGC-----------------GCTCCGATAATATTCCGCTAAATCCCAATA 4999
Db        931 GlnLeuAlaGlyMetLeuAlaGlyProAlaAlaAlaAlaSerIleLysSerProProVal 950

Qy       5000 GTGGCAATCGGCCGCC---------------------------------------------- 5015
Db        951 LeuGly-SerAlaAlaAlaSerProValHisLeuLysSerProSerLeuProAlaProSe 970

Qy       5016 ----------ACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAA 5065
Db        970 rProGlyTrpThrSerSerProGluProProLeuGlnSer---------------ProGl 985

Qy       5066 TGTCCCAACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTA 5125
Db        985 yIleProProAsnHisLysAlaProLeu-ThrMetAlaSerProAlaMetLeuGlyAsnV 1005

Qy       5126 TGACGATGATGGGGGGTCCG---------------GGACCGTCCGACATCAATATTGAGC 5170
Db       1005 alGluSerGlyGlyProProProProThrAlaSerGlnProAlaSerValAsnIleProG 1025

Qy       5171 ATGGAATAATTTCGGGACTAGATGGATCA------GGAATAGATACCATAAATCAAAATA 5224
Db       1025 lySerLeuProSerSerThrProTyrThrMetProProGluProThrLeuSerGlnAsnP 1045

Qy       5225 ACTGTCATTCAATGAATGTCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCT---- 5280
Db       1045 roLeuSerIleMetMetSerArgMetSerLysPheAlaMetProSerSerAsnProGlyT 1065

Qy       5281 ----------------AAAATGTGCGTAGCAGGCGGTCCAAATGGACCGCCTGGCTTTA 5323
Db       1065 yrAsnHisAspAlaIleLysThrValAlaSerSerAspAspAspSerProProAlaArgS 1085

Qy       5324 ATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGGCTGTGGCTCAG 5383
Db       1085 erProAsnLeuProSer---------MetAsnAsnMetProGlyMetGlyIleAsnThrG 1102

Qy       5384 CAAAC---------------------------------------------------------- 5388
Db       1102 lnAsnProArgIleSerGlyProAsnProValValProMetProThrLeuSerProMetG 1122

Qy       5389 -------------------------------------TCTTCAAACTTTCAAGGGGTT- 5409
Db       1122 lyMetThrGlnProLeuSerHisSerAsnGlnMetProSerProAsnAlaValGlyProA 1142

Qy       5410 ----------------GTTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTC---A 5449
Db       1142 snIleProProHisGlyValProMetGlyProGlyLeuMetSerHisAsnProIleMetG 1162

Qy       5450 ATTTTGGTTCGAATTTCAATCCGAATATT---CAGGTAAAGGCGAGTACCCCAAAC---- 5502
Db       1162 lyHisGlySerGlnGluProProMetValProGlnGlyArgMetGlyPheProGlnGlyP 1182

Qy       5503 -----ACCATACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATG 5557
Db       1182 heProProValGlnSerProProGlnGlnValProPheProHisAsnGlyProSerGlyG 1202

Qy       5558 GAGCTAATAATGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAACCCTC 5617
Db       1202 lyGlnGlySer------PheProGlyGlyMetGlyPhe-----------------ProG 1214

Qy       5618 AAATGGGTGCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGGTACTC 5677
Db       1214 lyGluGlyProLeuGlyArgProSerAsnLeu---ProGlnSerSerAlaAspAlaAlaL 1233

Qy       5678 CTGGAATGCCAGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTCT---- 5733
Db       1233 euCysLysProGlyGlyProGlyGlyProAspSerPheThrValLeuGlyAsnSerMetP 1253

Qy       5733 ---------------------------------------------------------- 5733
Db       1253 roSerValPheThrAspProAspLeuGlnGluValIleArgProGlyAlaThrGlyIleP 1273

Qy       5733 ---------------------------------------------------------- 5733
Db       1273 roGluPheAspLeuSerArgIleIleProSerGluLysProSerGlnThrLeuGlnTyrP 1293

Qy       5734 -----TCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTA 5788
Db       1293 heProArgGlyGluValProGlyArgLysGlnProGlnGlyProGlyPro------GlyP 1311

Qy       5789 TTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGT---------- 5838
Db       1311 heSerHisMetGln--------------GlyMetMetGlyGluGlnAlaProArgMetG 1326

Qy       5839 -----GGATTACCCGGCCATGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGG 5893
Db       1326 lyLeuAlaLeuProGlyMetGlyGlyProGlyProValGlyThrProAspIleProLeuG 1346

Qy       5894 GCATGCGTCCA--------------CATGCCATGCGGCCAAAT------GTAATGGGTG 5932
Db       1346 lyThrAlaProSerMetProGlyHisAsnProMetArgProProAlaPheLeuGlnGlnG 1366

Qy       5933 CGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACAGTCATCGGATGGTATTGACT 5992
Db       1366 lyMetMetGlyProHisHisArgMetMetSerProAlaGlnSerThr-----------M 1382

Qy       5993 GTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTGCAACAGCGCTGGACCACACA 6052
Db       1382 etProGlyGlnProThrLeuMetSerAsnProAlaAlaAla------------------- 1395

Qy       6053 TGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACAACACATAAAGAACATACCTA 6112
Db       1396 ----------------------------------------ValGlyMetIlePro- 1400

Qy       6113 GTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACAAGGGCAGATCCAACTGCATG 6172
Db       1401 --GlyLysAspArgGlyProAlaGlyLeuTyrThrHisProGlyProValGlySerProG 1420

Qy       6173 GGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTACTAATAATAAT 6222
Db       1420 lyMetMetMetSerMet---GlnGlyMetMetGlyProAsnArgThrSer 1435
```

RESULT 8
ADQ18945
ID ADQ18945 standard; protein; 1394 AA.
XX
AC ADQ18945;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 1764.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX

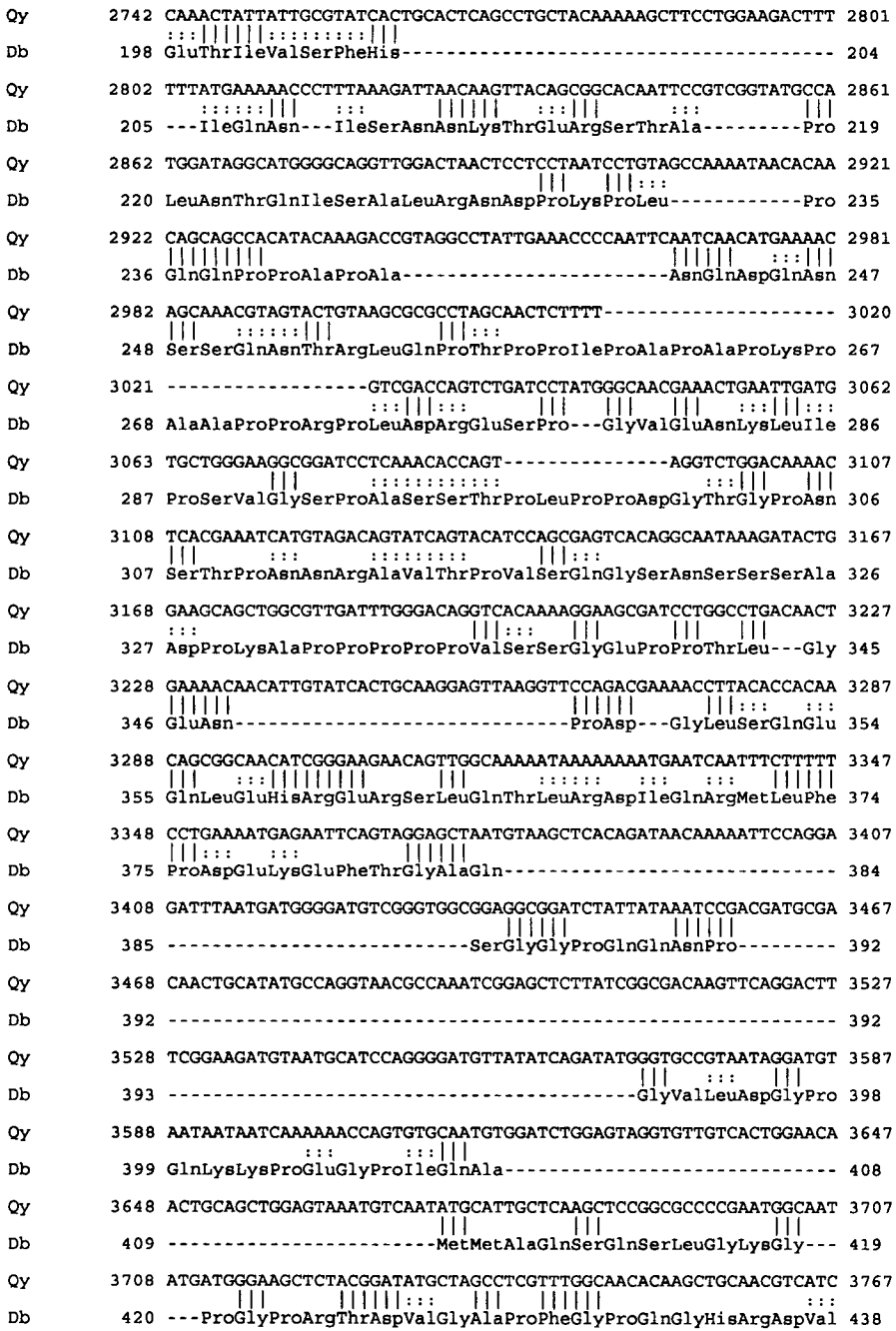
```
PS   Example 2; SEQ ID NO 1764; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   protein of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 1394 AA;

Alignment Scores:
Pred. No.:              3.12e-17     Length:        1394
Score:                  314.00       Matches:       306
Percent Similarity:     30.12%       Conservative:  172
Best Local Similarity:  19.28%       Mismatches:    509
Query Match:            2.61%        Indels:        601
DB:                     8            Gaps:          70

US-10-664-859-1 (1-6909) x ADQ18945 (1-1394)

Qy      2397 AAAATTGAAAGA---ATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACA 2453
              |||  ||||||  ||||||  ||||||        ::::::        :::
Db        98 LysArgGluArgSerIleSerAlaAspSerPheAspGlnArgAspProGlyThr------ 115

Qy      2454 ATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAAT 2513
                         ||||||:::      ::::::   ||||||      :::
Db       116 ---ProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHisIleLysSerGln 134

Qy      2514 GAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCAATACCC 2573
             :::|||   :::   |||         :::
Db       135 AspSerGlnHisThrProHisSerMetThr------------------------------ 144

Qy      2574 GGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCC 2633
                                     |||         ::::::         |||
Db       145 --------------------ProSerAsnAlaThrAlaProArgSerSerThrProSer 157

Qy      2634 TCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCAC------------ATA 2681
                |||   :::         :::         |||                     :::
Db       158 HisGlyGlnThrThrAlaThrGluProThrProAlaGlnLysThrProAlaLysValVal 177

Qy      2682 TTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTT 2741
             :::||||||||||||::::::||||||||   ||||||:::||||||    ||||||
Db       178 TyrValPheSerThrGluMetAlaAsnLysAlaAlaGluAlaValLeuLysGlyGlnVal 197

Qy      2742 CAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTT 2801
             :::||||||:::::::::|||
Db       198 GluThrIleValSerPheHis--------------------------------------- 204

Qy      2802 TTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCA 2861
                ::::::|||   :::      ||||||   :::|||      :::         |||
Db       205 ---IleGlnAsn---IleSerAsnAsnLysThrGluArgSerThrAla---------Pro 219

Qy      2862 TGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAA 2921
                                                |||   |||:::
Db       220 LeuAsnThrGlnIleSerAlaLeuArgAsnAspProLysProLeu------------Pro 235

Qy      2922 CAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAAC 2981
             |||||||||                                   ||||||   :::|||
Db       236 GlnGlnProProValProAla-----------------------AsnGlnAspGlnAsn 247

Qy      2982 AGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT-------------------- 3020
             |||   ::::::|||          |||:::
Db       248 SerSerGlnAsnThrArgLeuGlnProThrProProIleProAlaProAlaProLysPro 267

Qy      3021 ------------------GTCGACCAGTCTGATCCTATGGGCAACGAAACTGAATTGATG 3062
                         :::|||:::         |||    |||    |||    :::|||:::
Db       268 AlaAlaProProArgProLeuAspArgGluSerPro---GlyValGluAsnLysLeuIle 286

Qy      3063 TGCTGGGAAGGCGGATCCTCAAACACCAGT---------------AGGTCTGGACAAAAC 3107
                      |||      ::::::::::::                 :::|||   |||
Db       287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306

Qy      3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
             |||      :::      :::::::::      |||:::
Db       307 SerThrProAsnAsnArgAlaValThrProValSerGlnGlySerAsnSerSerSerAla 326

Qy      3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
             :::                        |||:::   |||         |||   |||
Db       327 AspProLysAlaProProProProValSerSerGlyGluProProThrLeu---Gly 345

Qy      3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
             ||||||                              ||||||   |||:::   :::
Db       346 GluAsn-----------------------------ProAsp---GlyLeuSerGlnGlu 354

Qy      3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
             |||   :::|||||||||      |||   ::::::   :::   :::   ||||||
Db       355 GlnLeuGluHisArgGluArgSerLeuGlnThrLeuArgAspIleGlnArgMetLeuPhe 374

Qy      3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
             |||:::   :::      ||||||
Db       375 ProAspGluLysGluPheThrGlyAlaGln------------------------------ 384

Qy      3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
                                        ||||||      ||||||
Db       385 --------------------------SerGlyGlyProGlnGlnAsnPro--------- 392

Qy      3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527

Db       392 ------------------------------------------------------------ 392

Qy      3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
                                                      |||   :::   |||
Db       393 ------------------------------------------GlyValLeuAspGlyPro 398

Qy      3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
                   :::      :::|||
Db       399 GlnLysLysProGluGlyProIleGlnAla----------------------------- 408

Qy      3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
                                    |||      |||            |||
Db       409 -----------------------MetMetAlaGlnSerGlnSerLeuGlyLysGly--- 419

Qy      3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
                |||      ||||||:::   |||   ||||||                  :::
Db       420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnGlyHisArgAspVal 438

Qy      3768 GGAACGGCCCCAGAT---ATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACC------ 3818
                :::||||||   |||         :::|||      |||   |||
Db       439 ProPheSerProAspGluMetValProProSerMetAsnSerGlnSerGlyThrIleGly 458

Qy      3819 ------CATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
                   |||                  |||:::   |||   |||:::|||   :::
Db       459 ProAspHisLeuAspHisMetThrProGluGlnIleAlaTrpLeuLysLeuGlnGlnGlu 478

Qy      3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
             |||:::||||||:::         ||||||   ||||||         ||||||
Db       479 PheTyrGluGluLysArgArg------LysProGluGlnVal---------ValValGln 493

Qy      3933 CAA--------------------------CAGCAAACCCCTTCTGGATCTGGTGGAAAC 3965
             |||                           |||   |||   |||
Db       494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgGly------------ 509

Qy      3966 TCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCAC--------- 4016
                                ::::::||||||||||||||||||
Db       510 --------------------ValValArgGlyProProProProTyrGlnMetThrPro 522

Qy      4017 ------------------------------------------TCCATCCAGAGATCTGCG 4034
                                                     |||
```

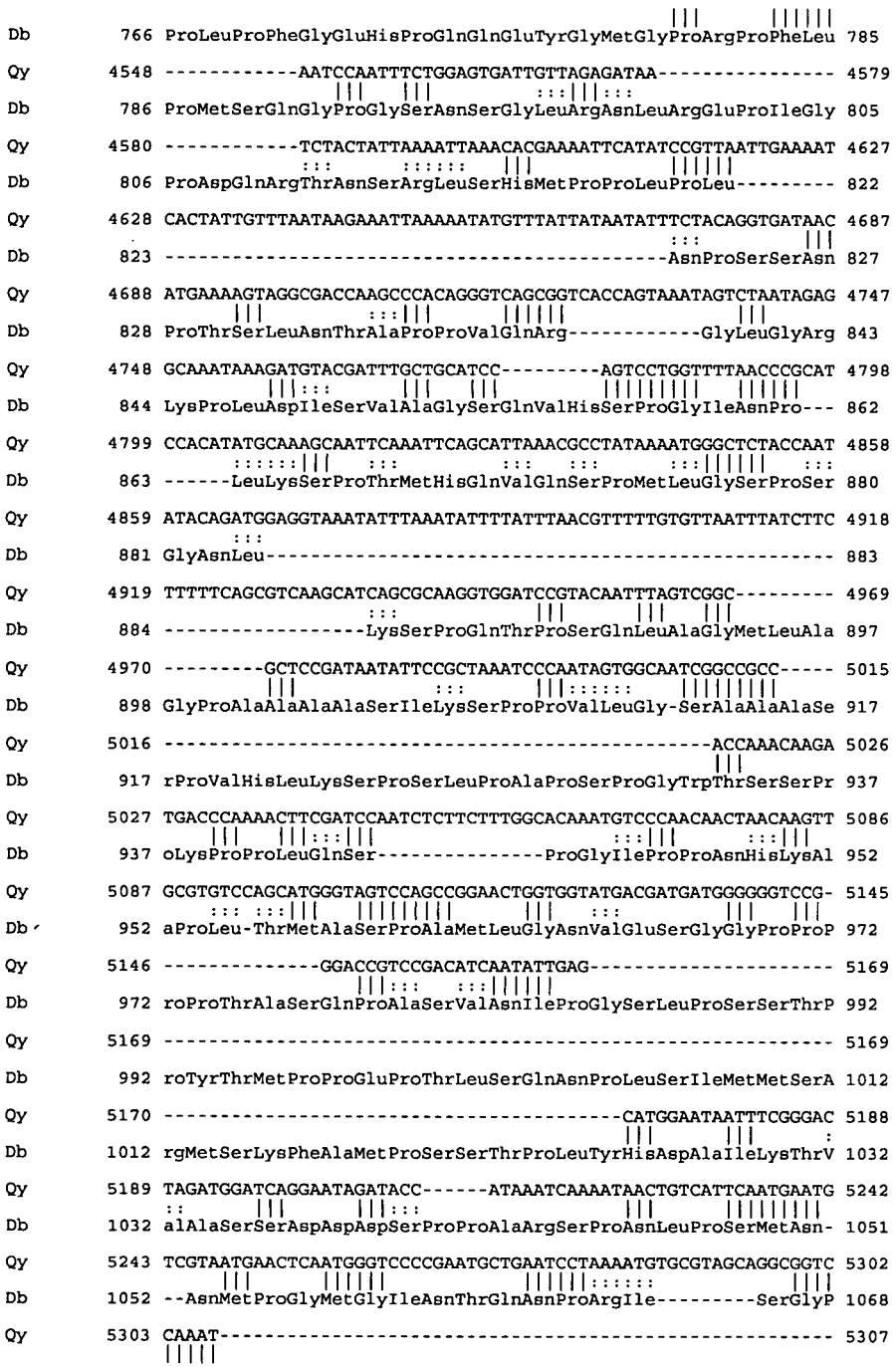

Db 523 SerGluGlyTrpAlaProGlyGlyThrGluProPheSerAspGlyIleAsnMetProHis 542
Qy 4035 TCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTA--------- 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProGlySerGlnMetArgLeu 562
Qy 4086 ----------TCTCTCCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGluMetGluGlyProAsnValProAsnProAla 582
Qy 4104 ACAACCGCAGCAGTCATGGGATTG------CCGACCAACTCTCCTAGCATG--------- 4148
Db 583 SerArgProGlyLeuSerGlyValSerTrpProAspAspValProLysIleProAspGly 602
Qy 4149 ------------GATGGAACAGGATCATTATCTGGA---------- 4172
Db 603 ArgAsnPheProProGlyArgGlyIlePheSerGlyProGlyArgGlyGluArgPhePro 622
Qy 4173 ----------TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGC 4211
Db 623 AsnProGlnGlyLeuSerGluGluMetPheGlnGlnGlnLeuAlaGluLysGlnLeuGly 642
Qy 4212 ACAACA------ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCG 4265
Db 643 LeuProProGlyMetAlaMetGluGlyIleArgProSerMetGluMetAsnArgMetIle 662
Qy 4266 CCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTA 4325
Db 663 ProGlySerGln---ArgHisMetGluProGlyAsnAsnProIlePheProArgIlePro 681
Qy 4326 AGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG---------- 4367
Db 682 ValGluGlyProLeuSerProSerArgGlyAspPheProLysGlyIleProProGlnMet 701
Qy 4368 ------------GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTT 4415
Db 702 GlyProGlyArgGluLeuGlu---PheGlyMetValProSerGlyMet---------- 716
Qy 4416 AATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAA 4475
Db 717 ---------LysGlyAspValAsnLeuAsn---ValAsn---------- 725
Qy 4476 CTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG--- 4532
Db 726 MetGlySerAsnSerGlnMetIleProGlnLysMetArgGluAlaGlyAlaGlyProGlu 745
Qy 4532 ---------- 4532
Db 746 GluMetLeuLysLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnLysMetVal 765
Qy 4533 ----------CCCATCCATTTCTTG 4547
Db 766 ProLeuProPheGlyGluHisProGlnGlnGluTyrGlyMetGlyProArgProPheLeu 785
Qy 4548 ------------AATCCAATTTCTGGAGTGATTGTTAGAGATAA---------- 4579
Db 786 ProMetSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
Qy 4580 ------------TCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAAT 4627
Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProProLeuProLeu--------- 822
Qy 4628 CACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTACAGGTGATAAC 4687
Db 823 ----------AsnProSerSerAsn 827
Qy 4688 ATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAG 4747
Db 828 ProThrSerLeuAsnThrAlaProProValGlnArg------------GlyLeuGlyArg 843
Qy 4748 GCAAATAAAGATGTACGATTTGCTGCATCC---------AGTCCTGGTTTTAACCCGCAT 4798
Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro--- 862
Qy 4799 CCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAAT 4858
Db 863 ------LeuLysSerProThrMetHisGlnValGlnSerProMetLeuGlySerProSer 880
Qy 4859 ATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTC 4918
Db 881 GlyAsnLeu---------- 883
Qy 4919 TTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGC--------- 4969
Db 884 ----------LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
Qy 4970 ---------GCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCC----- 5015
Db 898 GlyProAlaAlaAlaAlaSerIleLysSerProProValLeuGly-SerAlaAlaAlaSe 917
Qy 5016 ----------ACCAAACAAGA 5026
Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTrpThrSerSerPr 937
Qy 5027 TGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTT 5086
Db 937 oLysProProLeuGlnSer----------ProGlyIleProProAsnHisLysAl 952
Qy 5087 GCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCG- 5145
Db 952 aProLeu-ThrMetAlaSerProAlaMetLeuGlyAsnValGluSerGlyGlyProProP 972
Qy 5146 ----------GGACCGTCCGACATCAATATTGAG---------- 5169
Db 972 roProThrAlaSerGlnProAlaSerValAsnIleProGlySerLeuProSerSerThrP 992
Qy 5169 ---------- 5169
Db 992 roTyrThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012
Qy 5170 ----------CATGGAATAATTTCGGGAC 5188
Db 1012 rgMetSerLysPheAlaMetProSerSerThrProLeuTyrHisAspAlaIleLysThrV 1032
Qy 5189 TAGATGGATCAGGAATAGATACC------ATAAATCAAAATAACTGTCATTCAATGAATG 5242
Db 1032 alAlaSerSerAspAspAspSerProProAlaArgSerProAsnLeuProSerMetAsn- 1051
Qy 5243 TCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTC 5302
Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle---------SerGlyP 1068
Qy 5303 CAAAT---------- 5307
Db 1068 roAsnProValValProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088
Qy 5308 ----------GGACCGCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAA 5350
Db 1088 isSerAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107
Qy 5351 GAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTG 5410
Db 1108 --------ProMetGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
Qy 5411 TTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATC 5470
Db 1125 lnGluProProMetValProGlnGlyArgMet------GlyPheProGlnGlyPheProP 1143
Qy 5471 CGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATACATGCCAGTAAGGGCAC 5530
Db 1143 ro----------ValGlnSerProProGlnGlnValP 1152
Qy 5531 AGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGG 5590
Db 1152 roPheProHisAsnGlyProSerGlyGlyGlnGlySer------PheProGlyGlyMetG 1170

```
Qy       5591 AATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATAT 5650
Db       1170 lyPhe----------ProGlyGluGlyProLeuGlyArgProSerAsnLeu- 1183

Qy       5651 GCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAG 5710
Db       1184 --ProGlnSerSerAlaAspAlaAlaLeuCysLysProGlyGlyProGlyGlyProAspS 1203

Qy       5711 CCGGAGGTATGCTAATGAATTCT---------------------------------- 5733
Db       1203 erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV 1223

Qy       5733 ------------------------------------------------------------ 5733
Db       1223 alIleArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243

Qy       5734 --------------------------------TCCGGAGAGCAACACCAGAACAAGATCA 5761
Db       1243 luLysProSerGlnThrLeuGlnTyrPheProArgGlyGluValProGlyArgLysGlnP 1263

Qy       5762 CAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTA 5821
Db       1263 roGlnGlyProGlyPro------GlyPheSerHisMetGln----------GlyM 1276

Qy       5822 TTGTTGACGAAGAGGGT----------GGATTACCCGGCCATGACGGATCAATGA 5866
Db       1276 etMetGlyGluGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296

Qy       5867 ATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA----------CATGCCA 5911
Db       1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316

Qy       5912 TGCGGCCAAAT------GTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGT 5965
Db       1316 etArgProProAlaPheLeuGlnGlnGlyMetMetGlyProHisHisArgMetMetSerP 1336

Qy       5966 TTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACG 6025
Db       1336 roAlaGlnSerThr----------MetProGlyGlnProThrLeuMetSerAsnProA 1352

Qy       6026 CTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTA 6085
Db       1352 laAlaAla---------------------------------------------------- 1354

Qy       6086 AGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAG 6145
Db       1355 ----------ValGlyMetIlePro---GlyLysAspArgGlyProAlaGlyLeuTyrT 1370

Qy       6146 TGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTG 6205
Db       1370 hrHisProGlyProValGlySerProGlyMetMetMetSerMet---GlnGlyMetMetG 1389

Qy       6206 GACCTACTAATAATAAT 6222
Db       1389 lyProAsnArgThrSer 1394
```

```
RESULT 9
AAB71229
ID   AAB71229 standard; protein; 1426 AA.
XX
AC   AAB71229;
XX
DT   18-NOV-2002  (first entry)
XX
DE   Human legless homologue lgs/bc19 protein.
XX
KW   Legless; human; lgs; Wnt/Wingless signaling pathway; Wnt; Wg;
KW   tissue proliferation; tumour; cytostatic; cellular disorder; colon;
KW   blood disorder; cancer; breast; head and neck cancer; brain; thyroid;
KW   medulloblastoma; skin cancer; tissue regeneration; tissue repair.
XX
OS   Homo sapiens.
XX
PN   US2002086986-A1.
XX
PD   04-JUL-2002.
XX
PF   27-JUL-2001; 2001US-00915543.
XX
PR   28-JUL-2000; 2000US-0221502P.
XX
PA   (BASL/) BASLER K.
PA   (BRUN/) BRUNNER E.
PA   (FROE/) FROESCH B.
PA   (KRAM/) KRAMPS T.
PA   (PETE/) PETER O.
XX
PI   Basler K,  Brunner E,  Froesch B,  Kramps T,  Peter O;
XX
DR   WPI; 2002-635689/68.
DR   N-PSDB; AAF88467.
XX
PT   Novel polypeptide useful in therapeutic method for treating disorders of
PT   cell fate such as cell differentiation or cell proliferation.
XX
PS   Example II; Fig 8B; 41pp; English.
XX
CC   This invention describes a novel polypeptide sharing one or more
CC   homologous amino acid domains with the legless (lgs) protein, a
CC   downstream component of the Wnt/Wingless (Wnt/Wg) signaling pathway
CC   involved in the formation and maintenance of spatial arrangements and
CC   proliferation of tissues during development, and in the formation and
CC   growth of many human tumours. The products of the invention have
CC   cytostatic activity and can be used to treat cellular disorders, blood
CC   disorders and cancers caused by over-stimulation of the Wnt pathway,
CC   where the cancerous condition is colon, breast, head and neck, brain,
CC   thyroid, medulloblastoma or skin cancer. The product could also be used
CC   to promote tissue regeneration and repair. This sequence represents the
CC   human legless (lgs) protein homologue lgs/bcl9 described in the
CC   disclosure of the invention
XX
SQ   Sequence 1426 AA;
```

```
Alignment Scores:
Pred. No.:                4.69e-17      Length:          1426
Score:                    312.00        Matches:         313
Percent Similarity:       30.34%        Conservative:    177
Best Local Similarity:    19.38%        Mismatches:      519
Query Match:              2.59%         Indels:          607
DB:                       5             Gaps:            73

US-10-664-859-1 (1-6909) x AAB71229 (1-1426)

Qy       2397 AAAATTGAAAGA---ATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACA 2453
Db         98 LysArgGluArgSerIleSerAlaAspSerPheAspGlnArgAspProGlyThr------ 115

Qy       2454 ATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAAT 2513
Db        116 ---ProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHisIleLysSerGln 134

Qy       2514 GAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCAATACCC 2573
Db        135 AspSerGlnHisThrProHisSerMetThr------------------------------ 144

Qy       2574 GGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCC 2633
Db        145 ----------------------ProSerAsnAlaThrAlaProArgSerSerThrProSer 157

Qy       2634 TCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCAC------------ATA 2681
Db        158 HisGlyGlnThrThrAlaThrGluProThrProAlaGlnLysThrProAlaLysValVal 177

Qy       2682 TTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTT 2741
Db        178 TyrValPheSerThrGluMetAlaAsnLysAlaAlaGluAlaValLeuLysGlyGlnVal 197
```


Qy 2742 CAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTT 2801
Db 198 GluThrIleValSerPheHis 204
Qy 2802 TTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCA 2861
Db 205 ---IleGlnAsn---IleSerAsnAsnLysThrGluArgSerThrAla---------Pro 219
Qy 2862 TGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAA 2921
Db 220 LeuAsnThrGlnIleSerAlaLeuArgAsnAspProLysProLeu-----------Pro 235
Qy 2922 CAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAAC 2981
Db 236 GlnGlnProProAlaProAla----------AsnGlnAspGlnAsn 247
Qy 2982 AGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT---------- 3020
Db 248 SerSerGlnAsnThrArgLeuGlnProThrProProIleProAlaProAlaProLysPro 267
Qy 3021 ----------GTCGACCAGTCTGATCCTATGGGCAACGAAACTGAATTGATG 3062
Db 268 AlaAlaProProArgProLeuAspArgGluSerPro---GlyValGluAsnLysLeuIle 286
Qy 3063 TGCTGGGAAGGCGGATCCTCAAACACCAGT----------AGGTCTGGACAAAAC 3107
Db 287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306
Qy 3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
Db 307 SerThrProAsnAsnArgAlaValThrProValSerGlnGlySerAsnSerSerSerAla 326
Qy 3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
Db 327 AspProLysAlaProProProProProValSerSerGlyGluProProThrLeu---Gly 345
Qy 3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
Db 346 GluAsn----------ProAsp---GlyLeuSerGlnGlu 354
Qy 3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
Db 355 GlnLeuGluHisArgGluArgSerLeuGlnThrLeuArgAspIleGlnArgMetLeuPhe 374
Qy 3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
Db 375 ProAspGluLysGluPheThrGlyAlaGln---------- 384
Qy 3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
Db 385 ----------SerGlyGlyProGlnGlnAsnPro--------- 392
Qy 3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
Db 392 ---------- 392
Qy 3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
Db 393 ----------GlyValLeuAspGlyPro 398
Qy 3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
Db 399 GlnLysLysProGluGlyProIleGlnAla---------- 408
Qy 3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
Db 409 ----------MetMetAlaGlnSerGlnSerLeuGlyLysGly--- 419
Qy 3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
Db 420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnGlyHisArgAspVal 438
Qy 3768 GGAACGGCCCCAGAT---ATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACC------ 3818
Db 439 ProPheSerProAspGluMetValProProSerMetAsnSerGlnSerGlyThrIleGly 458
Qy 3819 ------CATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
Db 459 ProAspHisLeuAspHisMetThrProGluGlnIleAlaTrpLeuLysLeuGlnGlnGlu 478
Qy 3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
Db 479 PheTyrGluGluLysArgArg------LysGlnGluGlnVal---------ValValGln 493
Qy 3933 CAA----------CAGCAAACCCCTTCTGGATCTGGTGGAAAC 3965
Db 494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgGly---------- 509
Qy 3966 TCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCAC--------- 4016
Db 510 ----------ValValArgGlyProProProProTyrGlnMetThrPro 522
Qy 4017 ----------TCCATCCAGAGATCTGCG 4034
Db 523 SerGluGlyTrpAlaProGlyGlyThrGluProPheSerAspGlyIleAsnMetProHis 542
Qy 4035 TCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTA--------- 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProGlySerGlnMetArgLeu 562
Qy 4086 ----------TCTCTCCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGluMetGluGlyProAsnValProAsnProAla 582
Qy 4104 ACAACCGCAGCAGTCATGGGATTG------CCGACCAACTCTCCTAGCATG--------- 4148
Db 583 SerArgProGlyLeuSerGlyValSerTrpProAspAspValProLysIleProAspGly 602
Qy 4149 ----------GATGGAACAGGATCATTATCTGGA---------- 4172
Db 603 ArgAsnPheProProGlyGlnGlyIlePheSerGlyProGlyArgGlyGluArgPhePro 622
Qy 4173 ----------TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGC 4211
Db 623 AsnProGlnGlyLeuSerGluGluMetPheGlnGlnGlnLeuAlaGluLysGlnLeuGly 642
Qy 4212 ACAACA------ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCG 4265
Db 643 LeuProProGlyMetAlaMetGluGlyIleArgProSerMetGluMetAsnArgMetIle 662
Qy 4266 CCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTA 4325
Db 663 ProGlySerGln---ArgHisMetGluProGlyAsnAsnProIlePheProArgIlePro 681
Qy 4326 AGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG---------- 4367
Db 682 ValGluGlyProLeuSerProSerArgGlyAspPheProLysGlyIleProProGlnMet 701
Qy 4368 ----------GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTT 4415
Db 702 GlyProGlyArgGluLeuGlu---PheGlyMetValProSerGlyMet---------- 716
Qy 4416 AATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAA 4475
Db 717 ---------LysGlyAspValAsnLeuAsn---ValAsn---------- 725
Qy 4476 CTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG--- 4532
Db 726 MetGlySerAsnSerGlnMetIleProGlnLysMetArgGluAlaGlyAlaGlyProGlu 745
Qy 4532 ---------- 4532
Db 746 GluMetLeuLysLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnLysMetVal 765
Qy 4533 ----------CCCATCCATTTCTTG 4547


Db 766 ProLeuProPheGlyGluHisProGlnGlnGluTyrGlyMetGlyProArgProPheLeu 785
Qy 4548 ----------AATCCAATTTCTGGAGTGATTGTTAGAGATAA---------- 4579
Db 786 ProMetSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
Qy 4580 ----------TCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAAT 4627
Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProProLeuProLeu--------- 822
Qy 4628 CACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTACAGGTGATAAC 4687
Db 823 ----------AsnProSerSerAsn 827
Qy 4688 ATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAG 4747
Db 828 ProThrSerLeuAsnThrAlaProProValGlnArg----------GlyLeuGlyArg 843
Qy 4748 GCAAATAAAGATGTACGATTTGCTGCATCC---------AGTCCTGGTTTTAACCCGCAT 4798
Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro--- 862
Qy 4799 CCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAAT 4858
Db 863 ------LeuLysSerProThrMetHisGlnValGlnSerProMetLeuGlySerProSer 880
Qy 4859 ATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTC 4918
Db 881 GlyAsnLeu---------- 883
Qy 4919 TTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGC--------- 4969
Db 884 ----------LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
Qy 4970 ---------GCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCC----- 5015
Db 898 GlyProAlaAlaAlaAlaSerIleLysSerProProValLeuGly-SerAlaAlaAlaSe 917
Qy 5016 ----------ACCAAACAAGA 5026
Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTrpThrSerSerPr 937
Qy 5027 TGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTT 5086
Db 937 oLysProProLeuGlnSer----------ProGlyIleProProAsnHisLysAl 952
Qy 5087 GCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCG- 5145
Db 952 aProLeu-ThrMetAlaSerProAlaMetLeuGlyAsnValGluSerGlyGlyProProP 972
Qy 5146 ----------GGACCGTCCGACATCAATATTGAG---------- 5169
Db 972 roProThrAlaSerGlnProAlaSerValAsnIleProGlySerLeuProSerSerThrP 992
Qy 5169 ---------- 5169
Db 992 roTyrThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012
Qy 5170 ----------CATGGAATAATTTCGGGAC 5188
Db 1012 rgMetSerLysPheAlaMetProSerSerThrProLeuTyrHisAspAlaIleLysThrV 1032
Qy 5189 TAGATGGATCAGGAATAGATACC------ATAAATCAAAATAACTGTCATTCAATGAATG 5242
Db 1032 alAlaSerSerAspAspAspSerProProAlaArgSerProAsnLeuProSerMetAsn- 1051
Qy 5243 TCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTC 5302
Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle---------SerGlyP 1068
Qy 5303 CAAAT---------- 5307
Db 1068 roAsnProValValProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088
Qy 5308 ----------GGACCGCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAA 5350
Db 1088 isSerAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107
Qy 5351 GAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTG 5410
Db 1108 --------ProMetGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
Qy 5411 TTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATC 5470
Db 1125 lnGluProProMetValProGlnGlyArgMet------GlyPheProGlnGlyPheProP 1143
Qy 5471 CGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATACATGCCAGTAAGGGCAC 5530
Db 1143 ro----------ValGlnSerProProGlnGlnValP 1152
Qy 5531 AGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGG 5590
Db 1152 roPheProHisAsnGlyProSerGlyGlyGlnGlySer------PheProGlyGlyMetG 1170
Qy 5591 AATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATAT 5650
Db 1170 lyPhe----------ProGlyGluGlyProLeuGlyArgProSerAsnLeu- 1183
Qy 5651 GCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAG 5710
Db 1184 --ProGlnSerSerAlaAspAlaAlaLeuCysLysProGlyGlyProGlyGlyProAspS 1203
Qy 5711 CCGGAGGTATGCTAATGAATTCT---------- 5733
Db 1203 erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV 1223
Qy 5733 ---------- 5733
Db 1223 alIleArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243
Qy 5734 ----------TCCGGAGAGCAACACCAGAACAAGATCA 5761
Db 1243 luLysProSerGlnThrLeuGlnTyrPheProArgGlyGluValProGlyArgLysGlnP 1263
Qy 5762 CAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTA 5821
Db 1263 roGlnGlyProGlyPro------GlyPheSerHisMetGln----------GlyM 1276
Qy 5822 TTGTTGACGAAGAGGGT----------GGATTACCCGGCCATGACGGATCAATGA 5866
Db 1276 etMetGlyGluGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296
Qy 5867 ATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA----------CATGCCA 5911
Db 1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316
Qy 5912 TGCGGCCAAAT------GTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGT 5965
Db 1316 etArgProProAlaPheLeuGlnGlnGlyMetMetGlyProHisHisArgMetMetSerP 1336
Qy 5966 TTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACG 6025
Db 1336 roAlaGlnSerThr----------MetProGlyGlnProThrLeuMetSerAsnProA 1352
Qy 6026 CTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTA 6085
Db 1352 laAlaAla---------- 1354
Qy 6086 AGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAG 6145
Db 1355 ----------ValGlyMetIlePro---GlyLysAspArgGlyProAlaGlyLeuTyrT 1370
Qy 6146 TGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTG 6205
Db 1370 hrHisProGlyProValGlySerProGlyMetMetMetSerMet---GlnGlyMetMetG 1389

```
Qy       6206 GACCTACTAATAATAATTTA-----------ATGTCAACTGCCGGAAGTGTCAGTGCTA 6253
              |||||      |||   :::            |||         |||   ::::::|||
Db       1389 lyProGlnGlnAsnIleMetIleProProGlnMetArgProArgGly---MetAlaAlaA 1408

Qy       6254 CTAACGGTGTCTCTGGCATCAATTTCGTAGGTCCCTCTTCT 6294
                |||:::   |||   :::      ||||||   :::
Db       1408 spValGlyMetGlyGlyPheSer---GlnGlyProGlyAsn 1420
```

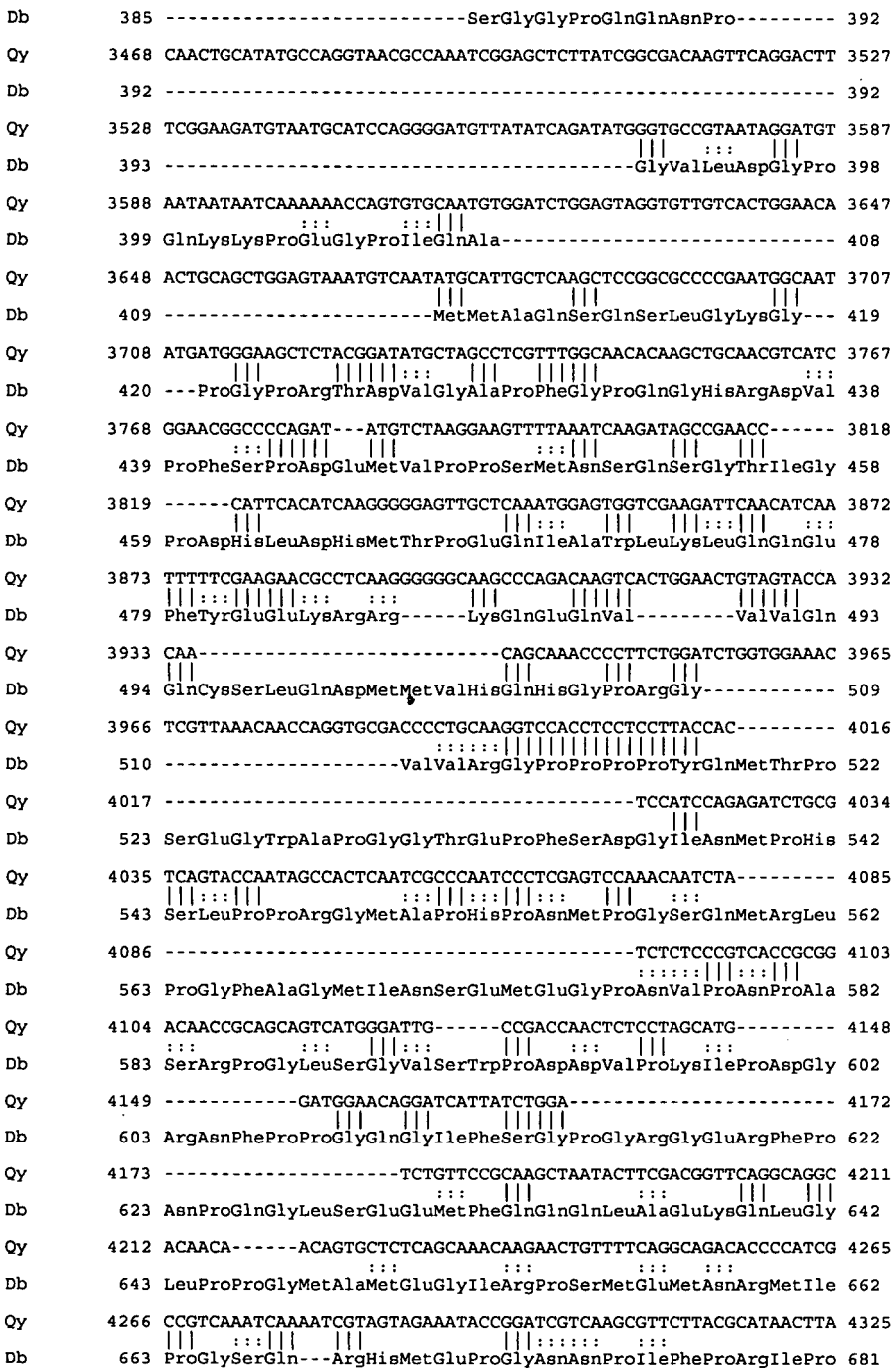
RESULT 10
ABW01534
ID ABW01534 standard; protein; 1426 AA.
XX
AC ABW01534;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human lgs/bcl9 protein.
XX
KW Legless protein; lgs; cell fate disorder; blood disease; gene therapy;
KW cancer; tissue regeneration; tissue repair; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003114413-A1.
XX
PD 19-JUN-2003.
XX
PF 19-DEC-2002; 2002US-00322579.
XX
PR 28-JUL-2000; 2000US-0221502P.
PR 27-JUL-2001; 2001US-00915543.
XX
PA (UYZU-) UNIV ZURICH.
XX
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;
XX
DR WPI; 2003-829432/77.
DR N-PSDB; AAD62642.
XX
PT Novel Lgs polypeptide useful for isolation of Lgs-binding proteins,
PT diagnosing disorders of cell fate, treating diseases such as cancer.
XX
PS Example 2; Fig 8B; 0pp; English.
XX
CC The invention relates to novel legless (lgs) proteins and polynucleotides
CC encoding such proteins. Lgs sequences are useful for the treatment of
CC disorders of cell fate such as differentiation or proliferation. The
CC invention is used to treat blood disease or a cancerous condition
CC characterised by over-stimulation of the Wnt pathway such as colon,
CC breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and
CC is administered to prevent progression from a pre-neoplastic or non-
CC malignant condition to a neoplastic or malignant state. It is
CC administered to promote tissue regeneration and repair. The invention is
CC also useful in the therapy of diseases cost by an over-activation of Wg
CC pathway. It is useful for reducing lgs gene expression in an invertebrate
CC or vertebrate organism or an invertebrate or vertebrate cell line. The
CC invention is also useful in gene therapy. The present sequence is human
CC lgs/bcl9 protein used in the invention
XX
SQ Sequence 1426 AA;

Alignment Scores:

Pred. No.:	4.69e-17	Length:	1426
Score:	312.00	Matches:	313
Percent Similarity:	30.34%	Conservative:	177
Best Local Similarity:	19.38%	Mismatches:	519
Query Match:	2.59%	Indels:	607
DB:	7	Gaps:	73

US-10-664-859-1 (1-6909) x ABW01534 (1-1426)

```
Qy       2397 AAAATTGAAAGA---ATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACA 2453
                |||   ||||||   ||||||   ||||||      ::::::         :::
Db         98 LysArgGluArgSerIleSerAlaAspSerPheAspGlnArgAspProGlyThr------ 115

Qy       2454 ATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAAT 2513
                  ||||||:::      ::::::   ||||||      :::
Db        116 ---ProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHisIleLysSerGln 134

Qy       2514 GAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCAATACCC 2573
                :::|||   :::   |||         :::
Db        135 AspSerGlnHisThrProHisSerMetThr---------------------------- 144

Qy       2574 GGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCC 2633
                                          |||                 ::::::      |||
Db        145 --------------------ProSerAsnAlaThrAlaProArgSerSerThrProSer 157

Qy       2634 TCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCAC------------ATA 2681
                |||   :::         :::          |||                      :::
Db        158 HisGlyGlnThrThrAlaThrGluProThrProAlaGlnLysThrProAlaLysValVal 177

Qy       2682 TTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTT 2741
                :::||||||||||||:::::||||||||   ||||||:::||||||   ||||||
Db        178 TyrValPheSerThrGluMetAlaAsnLysAlaAlaGluAlaValLeuLysGlyGlnVal 197

Qy       2742 CAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTT 2801
                :::||||||:::::::::|||
Db        198 GluThrIleValSerPheHis---------------------------------------- 204

Qy       2802 TTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCA 2861
                  ::::::|||   :::      ||||||   :::|||      :::         |||
Db        205 ---IleGlnAsn---IleSerAsnAsnLysThrGluArgSerThrAla---------Pro 219

Qy       2862 TGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAA 2921
                                              |||   |||:::
Db        220 LeuAsnThrGlnIleSerAlaLeuArgAsnAspProLysProLeu------------Pro 235

Qy       2922 CAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAAC 2981
                |||||||||                                     ||||||   :::|||
Db        236 GlnGlnProProAlaProAla-----------------------AsnGlnAspGlnAsn 247

Qy       2982 AGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT---------------------- 3020
                |||   ::::::|||         |||:::
Db        248 SerSerGlnAsnThrArgLeuGlnProThrProProIleProAlaProAlaProLysPro 267

Qy       3021 -----------------GTCGACCAGTCTGATCCTATGGGCAACGAAACTGAATTGATG 3062
                              :::|||:::      |||   |||   |||   :::|||:::
Db        268 AlaAlaProProArgProLeuAspArgGluSerPro---GlyValGluAsnLysLeuIle 286

Qy       3063 TGCTGGGAAGGCGGATCCTCAAACACCAGT--------------AGGTCTGGACAAAAC 3107
                  |||      ::::::::::::                  :::|||   |||
Db        287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306

Qy       3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
                |||      :::      :::::::::      |||:::
Db        307 SerThrProAsnAsnArgAlaValThrProValSerGlnGlySerAsnSerSerSerAla 326

Qy       3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
                :::                        |||:::   |||      |||   |||
Db        327 AspProLysAlaProProProProProValSerSerGlyGluProProThrLeu---Gly 345

Qy       3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
                ||||||                                ||||||      |||:::   :::
Db        346 GluAsn-----------------------------ProAsp---GlyLeuSerGlnGlu 354

Qy       3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAAATGAATCAATTTCTTTTT 3347
                |||   :::|||||||||      |||      ::::::   :::   :::   ||||||
Db        355 GlnLeuGluHisArgGluArgSerLeuGlnThrLeuArgAspIleGlnArgMetLeuPhe 374

Qy       3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
                |||:::   :::        ||||||
Db        375 ProAspGluLysGluPheThrGlyAlaGln---------------------------- 384

Qy       3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
                                          ||||||        ||||||
```


Db 385 ----------SerGlyGlyProGlnGlnAsnPro--------- 392
Qy 3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
Db 392 ---------- 392
Qy 3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
Db 393 ----------GlyValLeuAspGlyPro 398
Qy 3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
Db 399 GlnLysLysProGluGlyProIleGlnAla---------- 408
Qy 3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
Db 409 ----------MetMetAlaGlnSerGlnSerLeuGlyLysGly--- 419
Qy 3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
Db 420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnGlyHisArgAspVal 438
Qy 3768 GGAACGGCCCCAGAT---ATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACC------ 3818
Db 439 ProPheSerProAspGluMetValProProSerMetAsnSerGlnSerGlyThrIleGly 458
Qy 3819 ------CATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
Db 459 ProAspHisLeuAspHisMetThrProGluGlnIleAlaTrpLeuLysLeuGlnGlnGlu 478
Qy 3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
Db 479 PheTyrGluGluLysArgArg------LysGlnGluGlnVal---------ValValGln 493
Qy 3933 CAA----------CAGCAAACCCCTTCTGGATCTGGTGGAAAC 3965
Db 494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgGly---------- 509
Qy 3966 TCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCAC--------- 4016
Db 510 ----------ValValArgGlyProProProProTyrGlnMetThrPro 522
Qy 4017 ----------TCCATCCAGAGATCTGCG 4034
Db 523 SerGluGlyTrpAlaProGlyGlyThrGluProPheSerAspGlyIleAsnMetProHis 542
Qy 4035 TCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTA--------- 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProGlySerGlnMetArgLeu 562
Qy 4086 ----------TCTCTCCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGluMetGluGlyProAsnValProAsnProAla 582
Qy 4104 ACAACCGCAGCAGTCATGGGATTG------CCGACCAACTCTCCTAGCATG--------- 4148
Db 583 SerArgProGlyLeuSerGlyValSerTrpProAspAspValProLysIleProAspGly 602
Qy 4149 ----------GATGGAACAGGATCATTATCTGGA---------- 4172
Db 603 ArgAsnPheProProGlyGlnGlyIlePheSerGlyProGlyArgGlyGluArgPhePro 622
Qy 4173 ----------TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGC 4211
Db 623 AsnProGlnGlyLeuSerGluGluMetPheGlnGlnGlnLeuAlaGluLysGlnLeuGly 642
Qy 4212 ACAACA------ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCG 4265
Db 643 LeuProProGlyMetAlaMetGluGlyIleArgProSerMetGluMetAsnArgMetIle 662
Qy 4266 CCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTA 4325
Db 663 ProGlySerGln---ArgHisMetGluProGlyAsnAsnProIlePheProArgIlePro 681
Qy 4326 AGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG---------- 4367
Db 682 ValGluGlyProLeuSerProSerArgGlyAspPheProLysGlyIleProProGlnMet 701
Qy 4368 ----------GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTT 4415
Db 702 GlyProGlyArgGluLeuGlu---PheGlyMetValProSerGlyMet---------- 716
Qy 4416 AATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAA 4475
Db 717 ---------LysGlyAspValAsnLeuAsn---ValAsn---------- 725
Qy 4476 CTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG--- 4532
Db 726 MetGlySerAsnSerGlnMetIleProGlnLysMetArgGluAlaGlyAlaGlyProGlu 745
Qy 4532 ---------- 4532
Db 746 GluMetLeuLysLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnLysMetVal 765
Qy 4533 ----------CCCATCCATTTCTTG 4547
Db 766 ProLeuProPheGlyGluHisProGlnGlnGluTyrGlyMetGlyProArgProPheLeu 785
Qy 4548 ----------AATCCAATTTCTGGAGTGATTGTTAGAGATAA---------- 4579
Db 786 ProMetSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
Qy 4580 ----------TCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAAT 4627
Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProProLeuProLeu--------- 822
Qy 4628 CACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTACAGGTGATAAC 4687
Db 823 ----------AsnProSerSerAsn 827
Qy 4688 ATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAG 4747
Db 828 ProThrSerLeuAsnThrAlaProProValGlnArg----------GlyLeuGlyArg 843
Qy 4748 GCAAATAAAGATGTACGATTTGCTGCATCC---------AGTCCTGGTTTTAACCCGCAT 4798
Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro--- 862
Qy 4799 CCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAAT 4858
Db 863 ------LeuLysSerProThrMetHisGlnValGlnSerProMetLeuGlySerProSer 880
Qy 4859 ATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTC 4918
Db 881 GlyAsnLeu---------- 883
Qy 4919 TTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGC--------- 4969
Db 884 ----------LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
Qy 4970 ---------GCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCC----- 5015
Db 898 GlyProAlaAlaAlaAlaSerIleLysSerProProValLeuGly-SerAlaAlaAlaSe 917
Qy 5016 ----------ACCAAACAAGA 5026
Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTrpThrSerSerPr 937
Qy 5027 TGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTT 5086
Db 937 oLysProProLeuGlnSer----------ProGlyIleProProAsnHisLysAl 952
Qy 5087 GCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCG- 5145
Db 952 aProLeu-ThrMetAlaSerProAlaMetLeuGlyAsnValGluSerGlyGlyProProP 972

```
Qy      5146 -------------GGACCGTCCGACATCAATATTGAG---------------------- 5169
                          |||:::   :::||||||
Db       972 roProThrAlaSerGlnProAlaSerValAsnIleProGlySerLeuProSerSerThrP 992

Qy      5169 ------------------------------------------------------------ 5169

Db       992 roTyrThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012

Qy      5170 ----------------------------------------CATGGAATAATTTCGGGAC 5188
                                                     |||     |||      :
Db      1012 rgMetSerLysPheAlaMetProSerSerThrProLeuTyrHisAspAlaIleLysThrV 1032

Qy      5189 TAGATGGATCAGGAATAGATACC------ATAAATCAAAATAACTGTCATTCAATGAATG 5242
             ::      |||     |||:::          |||     |||||||||
Db      1032 alAlaSerSerAspAspAspSerProProAlaArgSerProAsnLeuProSerMetAsn- 1051

Qy      5243 TCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTC 5302
                |||     ||||||          |||||::::::          ||||
Db      1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle---------SerGlyP 1068

Qy      5303 CAAAT---------------------------------------------------- 5307
             |||||
Db      1068 roAsnProValValProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088

Qy      5308 ----------------GGACCGCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAA 5350
                            |||          ||||||   |||     |||:::
Db      1088 isSerAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107

Qy      5351 GAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTG 5410
                  :::|||   |||      |||   |||                 |||
Db      1108 --------ProMetGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125

Qy      5411 TTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATC 5470
                |||              |||||||||          |||      |||    |
Db      1125 lnGluProProMetValProGlnGlyArgMet------GlyPheProGlnGlyPheProP 1143

Qy      5471 CGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATACATGCCAGTAAGGGCAC 5530
             ||                                ::::|||      |||   :::
Db      1143 ro--------------------------------ValGlnSerProProGlnGlnValP 1152

Qy      5531 AGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGG 5590
                   :::|||      :::   |||      :::          |||      :::
Db      1152 roPheProHisAsnGlyProSerGlyGlyGlnGlySer------PheProGlyGlyMetG 1170

Qy      5591 AATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATAT 5650
               |||              |||      |||   :::|||      |||   :::
Db      1170 lyPhe-----------------ProGlyGluGlyProLeuGlyArgProSerAsnLeu- 1183

Qy      5651 GCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAG 5710
               |||   |||::::::|||          ||||||          ||||||   :
Db      1184 --ProGlnSerSerAlaAspAlaAlaLeuCysLysProGlyGlyProGlyGlyProAspS 1203

Qy      5711 CCGGAGGTATGCTAATGAATTCT----------------------------------- 5733
             ::      :::|||   ||||||
Db      1203 erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV 1223

Qy      5733 ------------------------------------------------------------ 5733

Db      1223 alIleArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243

Qy      5734 -----------------------------TCCGGAGAGCAACACCAGAACAAGATCA 5761
                                          ||||||          |||
Db      1243 luLysProSerGlnThrLeuGlnTyrPheProArgGlyGluValProGlyArgLysGlnP 1263

Qy      5762 CAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTA 5821
                |||||          |||   :::      |||                 :
Db      1263 roGlnGlyProGlyPro------GlyPheSerHisMetGln--------------GlyM 1276

Qy      5822 TTGTTGACGAAGAGGGT--------------GGATTACCCGGCCATGACGGATCAATGA 5866
             :::::   |||:::                 |||||||||      |||
Db      1276 etMetGlyGluGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296

Qy      5867 ATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA--------------CATGCCA 5911
```

```
                ::::|||   |||   :::      |||      |||                 :::   |
Db      1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316

Qy      5912 TGCGGCCAAAT------GTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGT 5965
             |||||||          :::                |||   |||   :::|||   :::
Db      1316 etArgProProAlaPheLeuGlnGlnGlyMetMetGlyProHisHisArgMetMetSerP 1336

Qy      5966 TTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACG 6025
               |||||||||:::                |||   |||:::                 |
Db      1336 roAlaGlnSerThr------------MetProGlyGlnProThrLeuMetSerAsnProA 1352

Qy      6026 CTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTA 6085
             ||:::
Db      1352 laAlaAla---------------------------------------------------- 1354

Qy      6086 AGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAG 6145
                        :::      ||||||   |||      :::      :::||||||
Db      1355 -----------ValGlyMetIlePro---GlyLysAspArgGlyProAlaGlyLeuTyrT 1370

Qy      6146 TGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTG 6205
                  |||   :::          |||          :::      |||   ::::::|
Db      1370 hrHisProGlyProValGlySerProGlyMetMetMetSerMet---GlnGlyMetMetG 1389

Qy      6206 GACCTACTAATAATAATTTA------------ATGTCAACTGCCGGAAGTGTCAGTGCTA 6253
             |||||      |||   :::            |||          |||   ::::::|||
Db      1389 lyProGlnGlnAsnIleMetIleProProGlnMetArgProArgGly---MetAlaAlaA 1408

Qy      6254 CTAACGGTGTCTCTGGCATCAATTTCGTAGGTCCCTCTTCT 6294
                |||:::   |||   :::      ||||||   :::
Db      1408 spValGlyMetGlyGlyPheSer---GlnGlyProGlyAsn 1420
```

RESULT 11
ADJ70152
ID ADJ70152 standard; protein; 1426 AA.
XX
AC ADJ70152;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID1958.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.

XX
PS Claim 1; SEQ ID NO 1958; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
SQ Sequence 1426 AA;

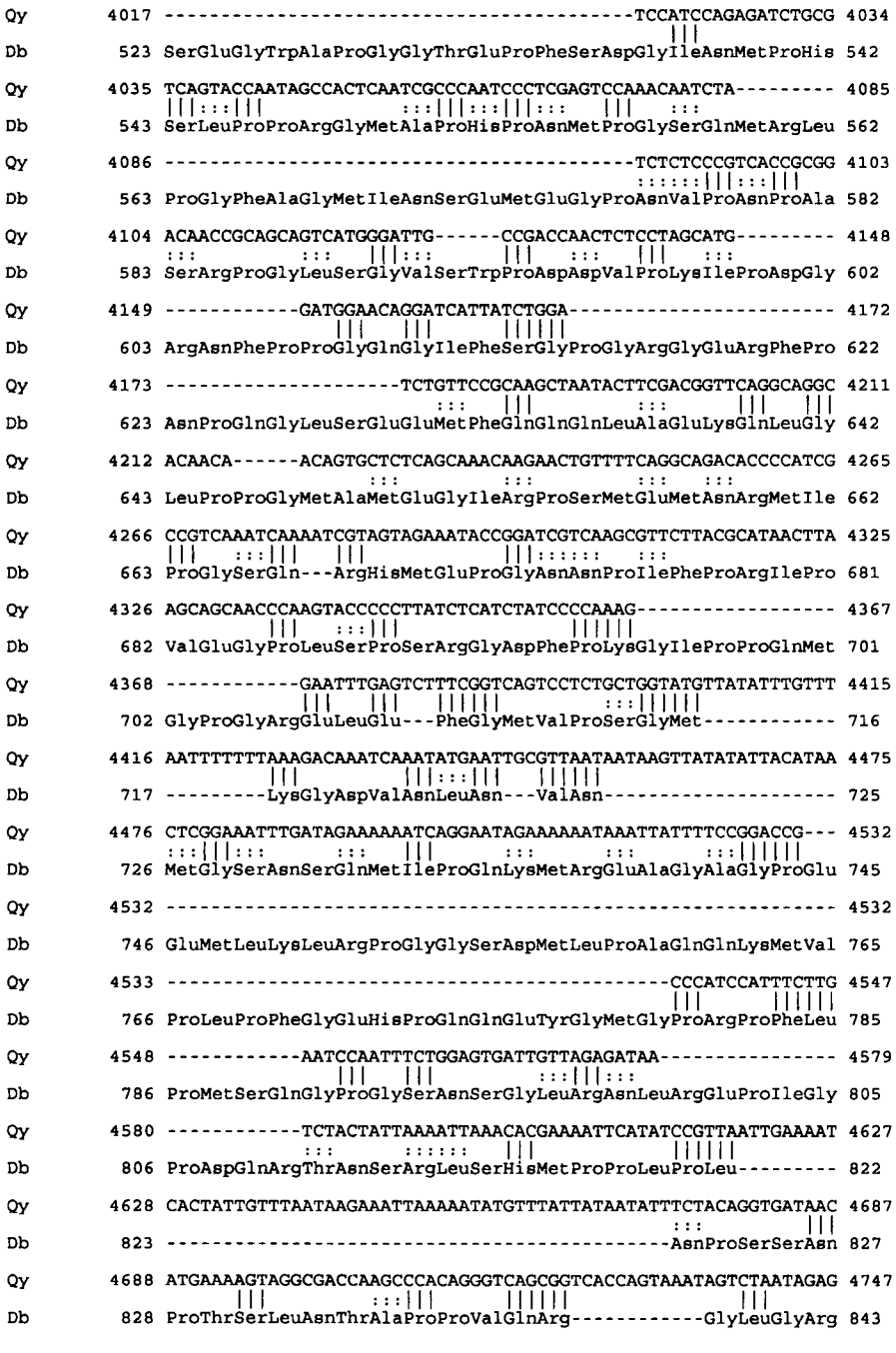
Alignment Scores:

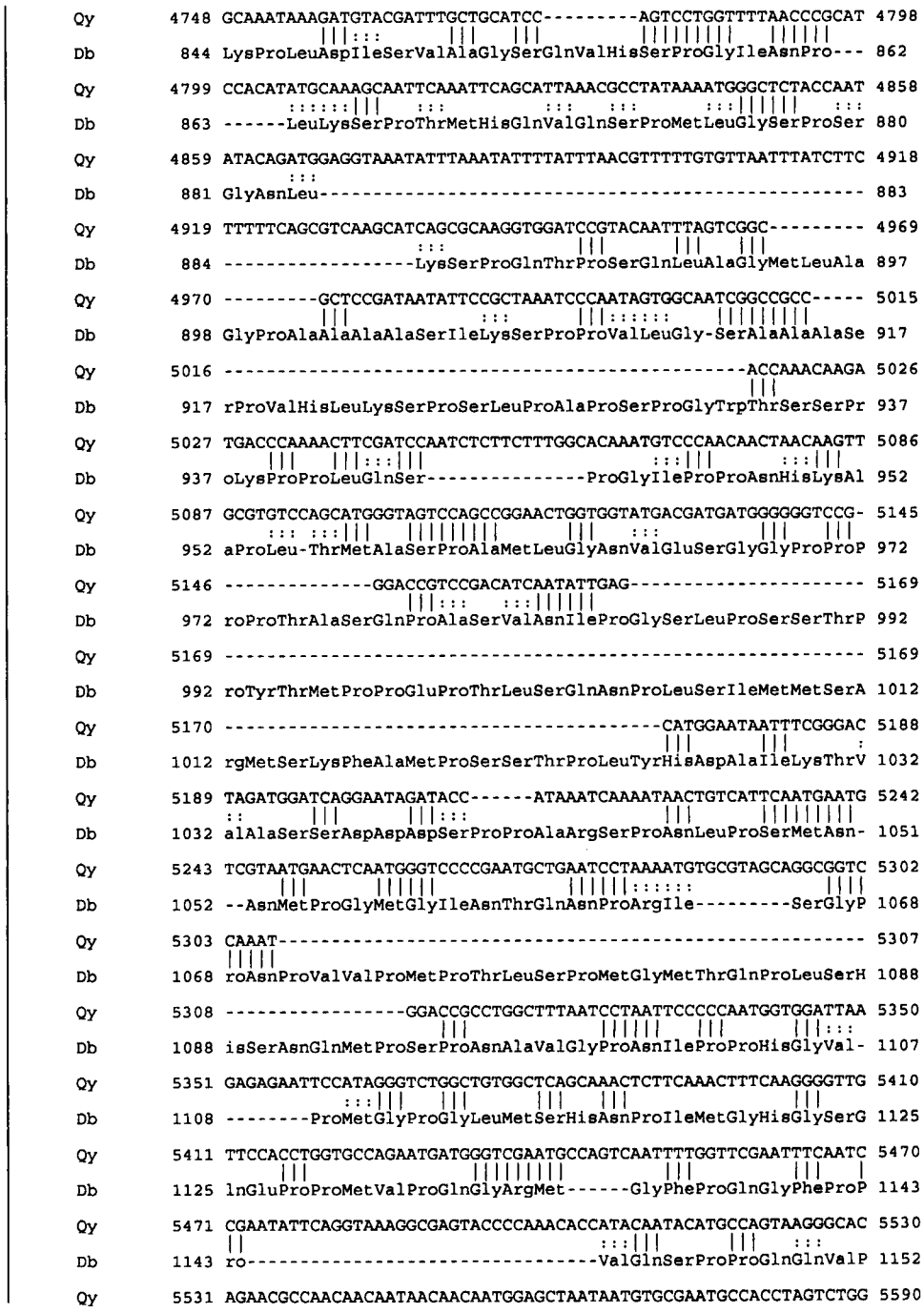
Pred. No.:	4.69e-17	Length:	1426
Score:	312.00	Matches:	313
Percent Similarity:	30.34%	Conservative:	177
Best Local Similarity:	19.38%	Mismatches:	519
Query Match:	2.59%	Indels:	607
DB:	7	Gaps:	73

US-10-664-859-1 (1-6909) x ADJ70152 (1-1426)

```
Qy       2397 AAAATTGAAAGA---ATTTCAAACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACA 2453
Db         98 LysArgGluArgSerIleSerAlaAspSerPheAspGlnArgAspProGlyThr------ 115

Qy       2454 ATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAAT 2513
Db        116 ---ProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHisIleLysSerGln 134

Qy       2514 GAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCAATACCC 2573
Db        135 AspSerGlnHisThrProHisSerMetThr---------------------------- 144

Qy       2574 GGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCC 2633
Db        145 ----------------------ProSerAsnAlaThrAlaProArgSerSerThrProSer 157

Qy       2634 TCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCAC------------ATA 2681
Db        158 HisGlyGlnThrThrAlaThrGluProThrProAlaGlnLysThrProAlaLysValVal 177

Qy       2682 TTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTT 2741
Db        178 TyrValPheSerThrGluMetAlaAsnLysAlaAlaGluAlaValLeuLysGlyGlnVal 197

Qy       2742 CAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTT 2801
Db        198 GluThrIleValSerPheHis------------------------------------- 204

Qy       2802 TTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCA 2861
Db        205 ---IleGlnAsn---IleSerAsnAsnLysThrGluArgSerThrAla---------Pro 219

Qy       2862 TGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAA 2921
Db        220 LeuAsnThrGlnIleSerAlaLeuArgAsnAspProLysProLeu------------Pro 235

Qy       2922 CAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAAC 2981
Db        236 GlnGlnProProAlaProAla-----------------------AsnGlnAspGlnAsn 247

Qy       2982 AGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT---------------------- 3020
```

```
Db        248 SerSerGlnAsnThrArgLeuGlnProThrProProIleProAlaProAlaProLysPro 267

Qy       3021 -----------------GTCGACCAGTCTGATCCTATGGGCAACGAAACTGAATTGATG 3062
Db        268 AlaAlaProProArgProLeuAspArgGluSerPro---GlyValGluAsnLysLeuIle 286

Qy       3063 TGCTGGGAAGGCGGATCCTCAAACACCAGT--------------AGGTCTGGACAAAAC 3107
Db        287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306

Qy       3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
Db        307 SerThrProAsnAsnArgAlaValThrProValSerGlnGlySerAsnSerSerSerAla 326

Qy       3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
Db        327 AspProLysAlaProProProProProValSerSerGlyGluProProThrLeu---Gly 345

Qy       3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
Db        346 GluAsn-----------------------------ProAsp---GlyLeuSerGlnGlu 354

Qy       3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
Db        355 GlnLeuGluHisArgGluArgSerLeuGlnThrLeuArgAspIleGlnArgMetLeuPhe 374

Qy       3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
Db        375 ProAspGluLysGluPheThrGlyAlaGln---------------------------- 384

Qy       3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
Db        385 --------------------------SerGlyGlyProGlnGlnAsnPro--------- 392

Qy       3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
Db        392 ------------------------------------------------------------ 392

Qy       3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
Db        393 ----------------------------------------GlyValLeuAspGlyPro 398

Qy       3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
Db        399 GlnLysLysProGluGlyProIleGlnAla---------------------------- 408

Qy       3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
Db        409 ------------------------MetMetAlaGlnSerGlnSerLeuGlyLysGly--- 419

Qy       3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
Db        420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnGlyHisArgAspVal 438

Qy       3768 GGAACGGCCCCAGAT---ATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACC------ 3818
Db        439 ProPheSerProAspGluMetValProProSerMetAsnSerGlnSerGlyThrIleGly 458

Qy       3819 ------CATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
Db        459 ProAspHisLeuAspHisMetThrProGluGlnIleAlaTrpLeuLysLeuGlnGlnGlu 478

Qy       3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
Db        479 PheTyrGluGluLysArgArg------LysGlnGluGlnVal---------ValValGln 493

Qy       3933 CAA-------------------------CAGCAAACCCCTTCTGGATCTGGTGGAAAC 3965
Db        494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgGly------------ 509

Qy       3966 TCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCAC--------- 4016
Db        510 --------------------ValValArgGlyProProProProTyrGlnMetThrPro 522
```

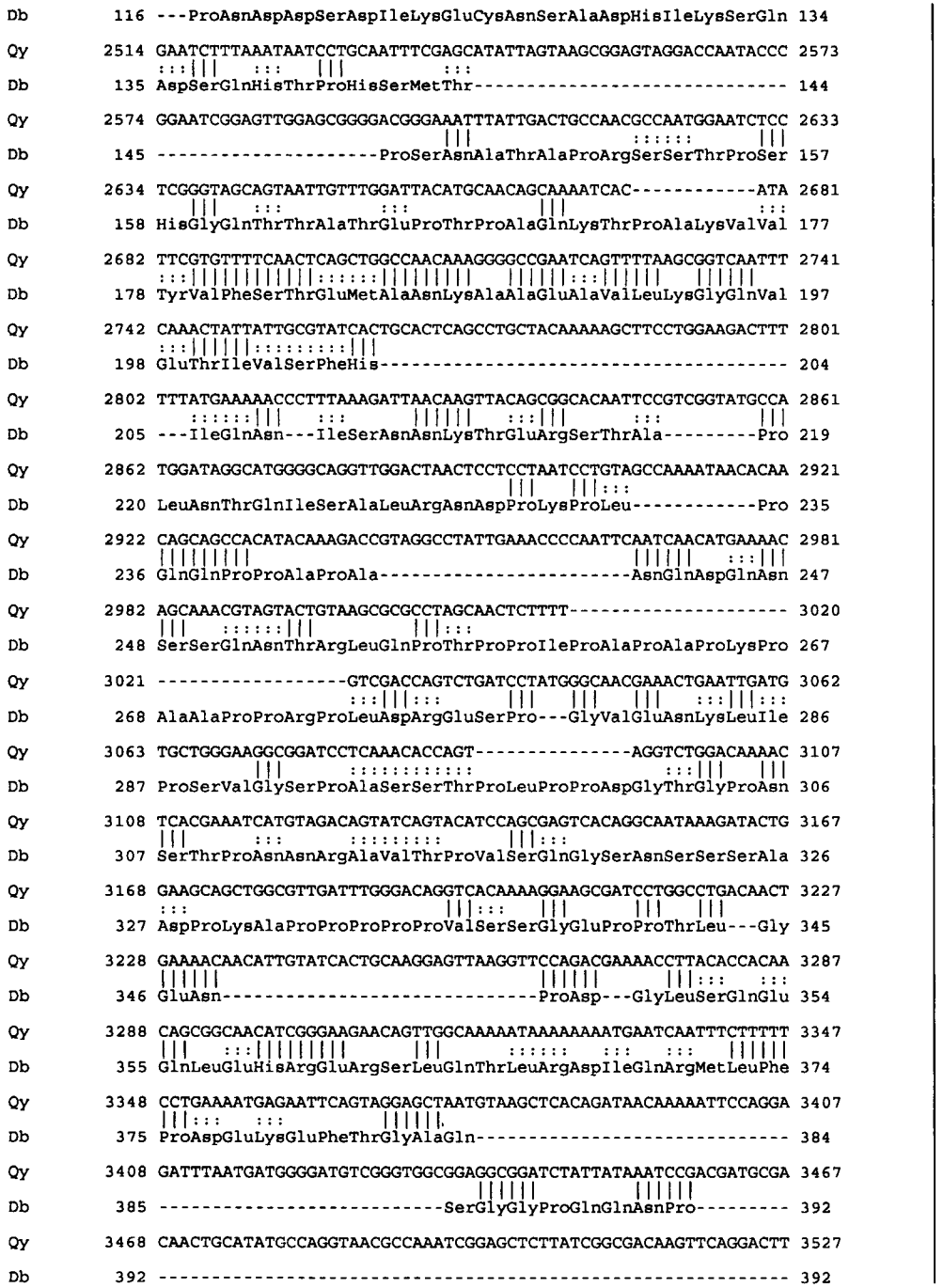

Qy 4017 TCCATCCAGAGATCTGCG 4034
Db 523 SerGluGlyTrpAlaProGlyGlyThrGluProPheSerAspGlyIleAsnMetProHis 542
Qy 4035 TCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTA 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProGlySerGlnMetArgLeu 562
Qy 4086 TCTCTCCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGluMetGluGlyProAsnValProAsnProAla 582
Qy 4104 ACAACCGCAGCAGTCATGGGATTG------CCGACCAACTCTCCTAGCATG 4148
Db 583 SerArgProGlyLeuSerGlyValSerTrpProAspAspValProLysIleProAspGly 602
Qy 4149 ------------GATGGAACAGGATCATTATCTGGA------------------------ 4172
Db 603 ArgAsnPheProProGlyGlnGlyIlePheSerGlyProGlyArgGlyGluArgPhePro 622
Qy 4173 --------------------TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGC 4211
Db 623 AsnProGlnGlyLeuSerGluGluMetPheGlnGlnGlnLeuAlaGluLysGlnLeuGly 642
Qy 4212 ACAACA------ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCG 4265
Db 643 LeuProProGlyMetAlaMetGluGlyIleArgProSerMetGluMetAsnArgMetIle 662
Qy 4266 CCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTA 4325
Db 663 ProGlySerGln---ArgHisMetGluProGlyAsnAsnProIlePheProArgIlePro 681
Qy 4326 AGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG------------------ 4367
Db 682 ValGluGlyProLeuSerProSerArgGlyAspPheProLysGlyIleProProGlnMet 701
Qy 4368 ------------GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTT 4415
Db 702 GlyProGlyArgGluLeuGlu---PheGlyMetValProSerGlyMet------------ 716
Qy 4416 AATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAA 4475
Db 717 ---------LysGlyAspValAsnLeuAsn---ValAsn---------- 725
Qy 4476 CTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG--- 4532
Db 726 MetGlySerAsnSerGlnMetIleProGlnLysMetArgGluAlaGlyAlaGlyProGlu 745
Qy 4532 ---------- 4532
Db 746 GluMetLeuLysLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnLysMetVal 765
Qy 4533 ----------CCCATCCATTTCTTG 4547
Db 766 ProLeuProPheGlyGluHisProGlnGlnGluTyrGlyMetGlyProArgProPheLeu 785
Qy 4548 ------------AATCCAATTTCTGGAGTGATTGTTAGAGATAA---------- 4579
Db 786 ProMetSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
Qy 4580 ------------TCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAAT 4627
Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProProLeuProLeu--------- 822
Qy 4628 CACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTACAGGTGATAAC 4687
Db 823 ----------AsnProSerSerAsn 827
Qy 4688 ATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAG 4747
Db 828 ProThrSerLeuAsnThrAlaProProValGlnArg------------GlyLeuGlyArg 843

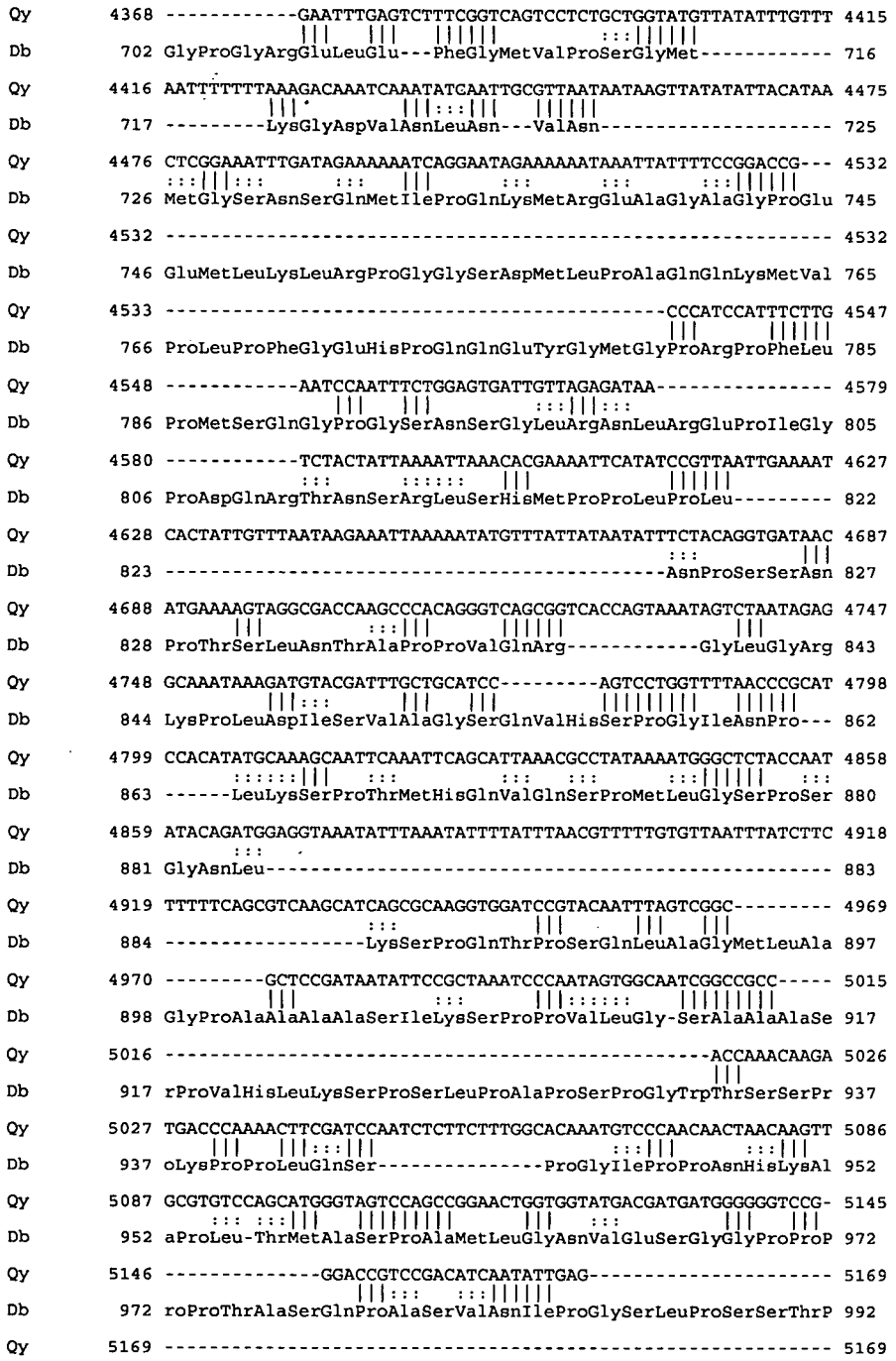

Qy 4748 GCAAATAAAGATGTACGATTTGCTGCATCC---------AGTCCTGGTTTTAACCCGCAT 4798
Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro--- 862
Qy 4799 CCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAAT 4858
Db 863 ------LeuLysSerProThrMetHisGlnValGlnSerProMetLeuGlySerProSer 880
Qy 4859 ATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTC 4918
Db 881 GlyAsnLeu---------- 883
Qy 4919 TTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGC--------- 4969
Db 884 ------------------LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
Qy 4970 ---------GCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCC----- 5015
Db 898 GlyProAlaAlaAlaAlaSerIleLysSerProProValLeuGly-SerAlaAlaAlaSe 917
Qy 5016 ----------ACCAAACAAGA 5026
Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTrpThrSerSerPr 937
Qy 5027 TGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTT 5086
Db 937 oLysProProLeuGlnSer---------------ProGlyIleProProAsnHisLysAl 952
Qy 5087 GCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCG- 5145
Db 952 aProLeu-ThrMetAlaSerProAlaMetLeuGlyAsnValGluSerGlyGlyProProP 972
Qy 5146 --------------GGACCGTCCGACATCAATATTGAG---------- 5169
Db 972 roProThrAlaSerGlnProAlaSerValAsnIleProGlySerLeuProSerSerThrP 992
Qy 5169 ---------- 5169
Db 992 roTyrThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012
Qy 5170 ----------CATGGAATAATTTCGGGAC 5188
Db 1012 rgMetSerLysPheAlaMetProSerSerThrProLeuTyrHisAspAlaIleLysThrV 1032
Qy 5189 TAGATGGATCAGGAATAGATACC------ATAAATCAAAATAACTGTCATTCAATGAATG 5242
Db 1032 alAlaSerSerAspAspSerProProAlaArgSerProAsnLeuProSerMetAsn- 1051
Qy 5243 TCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTC 5302
Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle---------SerGlyP 1068
Qy 5303 CAAAT---------- 5307
Db 1068 roAsnProValValProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088
Qy 5308 ----------GGACCGCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAA 5350
Db 1088 isSerAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107
Qy 5351 GAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTG 5410
Db 1108 --------ProMetGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
Qy 5411 TTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATC 5470
Db 1125 lnGluProProMetValProGlnGlyArgMet------GlyPheProGlnGlyPheProP 1143
Qy 5471 CGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATACATGCCAGTAAGGGCAC 5530
Db 1143 ro----------ValGlnSerProProGlnGlnValP 1152
Qy 5531 AGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGG 5590

```
                  :::|||      :::   |||      :::         |||      :::
Db       1152 roPheProHisAsnGlyProSerGlyGlyGlnGlySer------PheProGlyGlyMetG 1170

Qy       5591 AATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATAT 5650
               |||                  |||      |||  :::|||      |||   :::
Db       1170 lyPhe-----------------ProGlyGluGlyProLeuGlyArgProSerAsnLeu- 1183

Qy       5651 GCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAG 5710
                 |||   |||::::::|||               ||||||         ||||||   :
Db       1184 --ProGlnSerSerAlaAspAlaAlaLeuCysLysProGlyGlyProGlyGlyProAspS 1203

Qy       5711 CCGGAGGTATGCTAATGAATTCT---------------------------------- 5733
               ::      :::|||   ||||||
Db       1203 erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV 1223

Qy       5733 ------------------------------------------------------------ 5733

Db       1223 alIleArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243

Qy       5734 -------------------------------TCCGGAGAGCAACACCAGAACAAGATCA 5761
                                               ||||||           |||
Db       1243 luLysProSerGlnThrLeuGlnTyrPheProArgGlyGluValProGlyArgLysGlnP 1263

Qy       5762 CAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTA 5821
                 ||||||         |||   :::      |||                :
Db       1263 roGlnGlyProGlyPro------GlyPheSerHisMetGln--------------GlyM 1276

Qy       5822 TTGTTGACGAAGAGGGT---------------GGATTACCCGGCCATGACGGATCAATGA 5866
               :::::   |||:::                 |||||||||      |||
Db       1276 etMetGlyGluGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296

Qy       5867 ATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA---------------CATGCCA 5911
                 :::|||   |||   :::      |||      |||                :::   |
Db       1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316

Qy       5912 TGCGGCCAAAT------GTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGT 5965
               ||||||||         :::           |||   |||   :::|||   :::
Db       1316 etArgProProAlaPheLeuGlnGlnGlyMetMetGlyProHisHisArgMetMetSerP 1336

Qy       5966 TTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACG 6025
                 |||||||||:::                 |||   |||:::                |
Db       1336 roAlaGlnSerThr------------MetProGlyGlnProThrLeuMetSerAsnProA 1352

Qy       6026 CTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTA 6085
               ||:::
Db       1352 laAlaAla---------------------------------------------------- 1354

Qy       6086 AGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAG 6145
                     :::      ||||||   |||      :::      :::||||||
Db       1355 -----------ValGlyMetIlePro---GlyLysAspArgGlyProAlaGlyLeuTyrT 1370

Qy       6146 TGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTG 6205
                 |||   :::          |||          :::      |||   ::::::|
Db       1370 hrHisProGlyProValGlySerProGlyMetMetMetSerMet---GlnGlyMetMetG 1389

Qy       6206 GACCTACTAATAATAATTTA------------ATGTCAACTGCCGGAAGTGTCAGTGCTA 6253
               |||||      |||   :::            |||        |||   ::::::|||
Db       1389 lyProGlnGlnAsnIleMetIleProProGlnMetArgProArgGly---MetAlaAlaA 1408

Qy       6254 CTAACGGTGTCTCTGGCATCAATTTCGTAGGTCCCTCTTCT 6294
                   |||:::   |||   :::      ||||||   :::
Db       1408 spValGlyMetGlyGlyPheSer---GlnGlyProGlyAsn 1420
```

```
RESULT 12
ADJ71903
ID   ADJ71903 standard; protein; 1426 AA.
XX
AC   ADJ71903;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human Lgs/Bc19 polypeptide.
XX
KW   Human; legless; lgs; cell differentiation disorder;
KW   cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;
KW   breast; head; neck; brain; thyroid; skin; blood disease;
KW   tissue regeneration; tissue repair; cytostatic; Lgs/Bc19.
XX
OS   Homo sapiens.
XX
PN   US2004038901-A1.
XX
PD   26-FEB-2004.
XX
PF   22-SEP-2003; 2003US-00664859.
XX
PR   28-JUL-2000; 2000US-0221502P.
PR   27-JUL-2001; 2001US-00915543.
XX
PA   (UYZU-) UNIV ZURICH.
XX
PI   Basler K,  Brunner E,  Froesch B,  Kramps T,  Peter O;
XX
DR   WPI; 2004-203288/19.
DR   N-PSDB; ADJ71902.
XX
PT   Novel polypeptide sharing one or more homologue amino acid domains with
PT   Legless protein being functional homologue of Legless, useful for
PT   diagnosing disorders of cell fate.
XX
PS   Example 2; SEQ ID NO 15; 62pp; English.
XX
CC   The invention relates to a polypeptide sharing one or more homologous
CC   amino acid domains with a Legless (Lgs) protein and is therefore a
CC   functional homologue of Lgs. The invention also relates to a nucleotide
CC   sequence encoding a protein present in invertebrate and/or vertebrate
CC   organisms, the nucleotide sequence encoding a protein comprising a
CC   positive function in a regulatory pathway and the use of the polypeptide
CC   for the isolation of Lgs-binding proteins by carrying out an assay chosen
CC   from an in vitro binding assay with such a peptide or a co-
CC   immunoprecipitation from vertebrate or invertebrate cell lysates or a
CC   mammalian or yeast two hybrid assay. The polypeptide and polynucleotide
CC   are useful for treating disorders of cell fate, which involves
CC   administering therapeutic compounds chosen from invertebrate and
CC   vertebrate Lgs protein homologues or fragments, antibodies, antibody
CC   fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA,
CC   small peptides or chemical and natural compounds being capable of
CC   interfering with Lgs function, synthesis and degradation. The disorders
CC   are related to cell differentiation or cell proliferation. The compound
CC   is administered to treat a cancerous condition by preventing progression
CC   from a pre-neoplastic or non-malignant condition to a neoplastic or
CC   malignant state. The cancerous condition is characterised by over-
CC   stimulation of the Wnt pathway and is medulloblastoma or cancer of the
CC   colon, breast, head and neck, brain, thyroid or skin. The therapeutic
CC   compound may also be administered to a blood disease to promote tissue
CC   regeneration and repair. This sequence represents the human Lgs/Bc19
CC   polypeptide of the invention.
XX
SQ   Sequence 1426 AA;
```

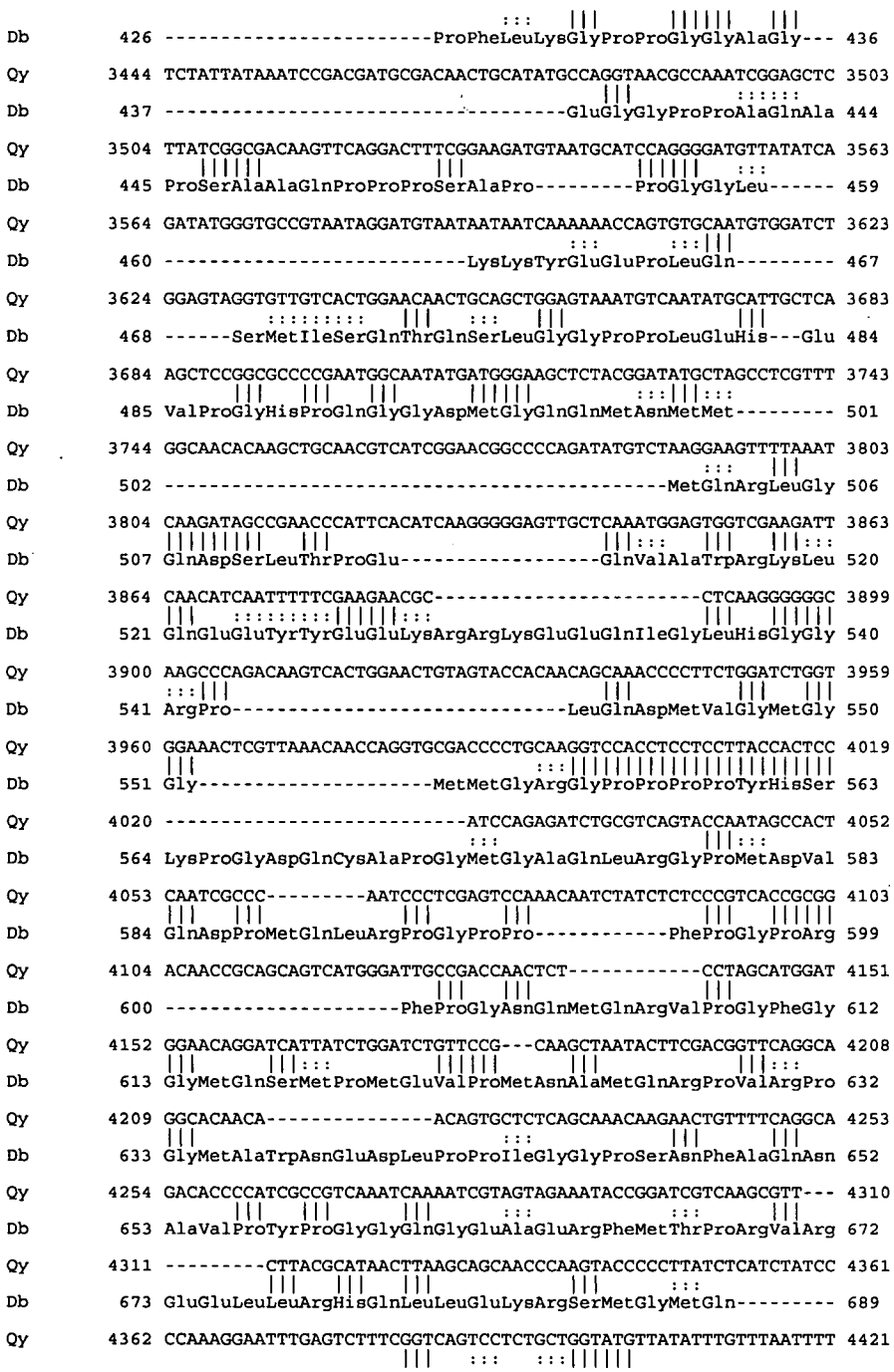
```
Alignment Scores:
Pred. No.:              4.69e-17        Length:         1426
Score:                  312.00          Matches:        313
Percent Similarity:     30.34%          Conservative:   177
Best Local Similarity:  19.38%          Mismatches:     519
Query Match:            2.59%           Indels:         607
DB:                     8               Gaps:           73

US-10-664-859-1 (1-6909) x ADJ71903 (1-1426)

Qy       2397 AAAATTGAAAGA---ATTTCAAACGACAGTACCACGGAAAAAAAAAGGATCGTCCTTGACA 2453
                 |||   ||||||   ||||||   ||||||       ::::::          :::
Db         98 LysArgGluArgSerIleSerAlaAspSerPheAspGlnArgAspProGlyThr------ 115

Qy       2454 ATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAAT 2513
                       ||||||:::      ::::::   ||||||      :::
```

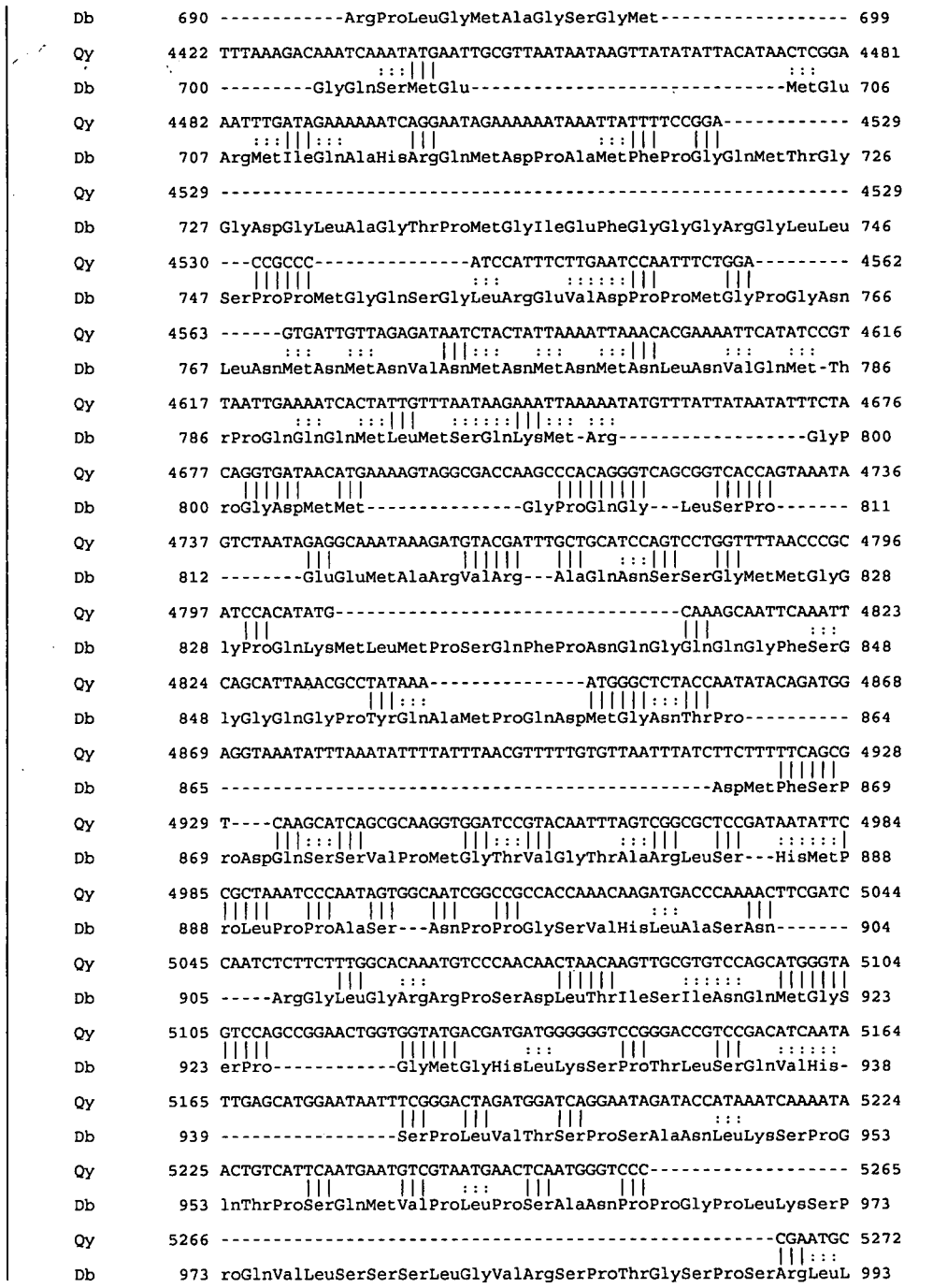

Db 116 ---ProAsnAspAspSerAspIleLysGluCysAsnSerAlaAspHisIleLysSerGln 134
Qy 2514 GAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTAAGCGGAGTAGGACCAATACCC 2573
Db 135 AspSerGlnHisThrProHisSerMetThr---------- 144
Qy 2574 GGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCC 2633
Db 145 ----------ProSerAsnAlaThrAlaProArgSerSerThrProSer 157
Qy 2634 TCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAAAATCAC----------ATA 2681
Db 158 HisGlyGlnThrThrAlaThrGluProThrProAlaGlnLysThrProAlaLysValVal 177
Qy 2682 TTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTT 2741
Db 178 TyrValPheSerThrGluMetAlaAsnLysAlaAlaGluAlaValLeuLysGlyGlnVal 197
Qy 2742 CAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTT 2801
Db 198 GluThrIleValSerPheHis---------- 204
Qy 2802 TTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCA 2861
Db 205 ---IleGlnAsn---IleSerAsnAsnLysThrGluArgSerThrAla---------Pro 219
Qy 2862 TGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAA 2921
Db 220 LeuAsnThrGlnIleSerAlaLeuArgAsnAspProLysProLeu----------Pro 235
Qy 2922 CAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAAC 2981
Db 236 GlnGlnProProAlaProAla----------AsnGlnAspGlnAsn 247
Qy 2982 AGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTT---------- 3020
Db 248 SerSerGlnAsnThrArgLeuGlnProThrProProIleProAlaProAlaProLysPro 267
Qy 3021 ----------GTCGACCAGTCTGATCCTATGGGCAACGAAACTGAATTGATG 3062
Db 268 AlaAlaProProArgProLeuAspArgGluSerPro---GlyValGluAsnLysLeuIle 286
Qy 3063 TGCTGGGAAGGCGGATCCTCAAACACCAGT----------AGGTCTGGACAAAAC 3107
Db 287 ProSerValGlySerProAlaSerSerThrProLeuProProAspGlyThrGlyProAsn 306
Qy 3108 TCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTG 3167
Db 307 SerThrProAsnAsnArgAlaValThrProValSerGlnGlySerAsnSerSerSerAla 326
Qy 3168 GAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACT 3227
Db 327 AspProLysAlaProProProProProValSerSerGlyGluProProThrLeu---Gly 345
Qy 3228 GAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAA 3287
Db 346 GluAsn----------ProAsp---GlyLeuSerGlnGlu 354
Qy 3288 CAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTT 3347
Db 355 GlnLeuGluHisArgGluArgSerLeuGlnThrLeuArgAspIleGlnArgMetLeuPhe 374
Qy 3348 CCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGA 3407
Db 375 ProAspGluLysGluPheThrGlyAlaGln---------- 384
Qy 3408 GATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGA 3467
Db 385 ----------SerGlyGlyProGlnGlnAsnPro--------- 392
Qy 3468 CAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTT 3527
Db 392 ---------- 392
Qy 3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
Db 393 ----------GlyValLeuAspGlyPro 398
Qy 3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACA 3647
Db 399 GlnLysLysProGluGlyProIleGlnAla---------- 408
Qy 3648 ACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCGAATGGCAAT 3707
Db 409 ----------MetMetAlaGlnSerGlnSerLeuGlyLysGly--- 419
Qy 3708 ATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATC 3767
Db 420 ---ProGlyProArgThrAspValGlyAlaProPheGlyProGlnGlyHisArgAspVal 438
Qy 3768 GGAACGGCCCCAGAT---ATGTCTAAGGAAGTTTTAAATCAAGATAGCCGAACC------ 3818
Db 439 ProPheSerProAspGluMetValProProSerMetAsnSerGlnSerGlyThrIleGly 458
Qy 3819 ------CATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
Db 459 ProAspHisLeuAspHisMetThrProGluGlnIleAlaTrpLeuLysLeuGlnGlnGlu 478
Qy 3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
Db 479 PheTyrGluGluLysArgArg------LysGlnGluGlnVal---------ValValGln 493
Qy 3933 CAA----------CAGCAAACCCCTTCTGGATCTGGTGGAAAC 3965
Db 494 GlnCysSerLeuGlnAspMetMetValHisGlnHisGlyProArgGly---------- 509
Qy 3966 TCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCAC--------- 4016
Db 510 ----------ValValArgGlyProProProProTyrGlnMetThrPro 522
Qy 4017 ----------TCCATCCAGAGATCTGCG 4034
Db 523 SerGluGlyTrpAlaProGlyGlyThrGluProPheSerAspGlyIleAsnMetProHis 542
Qy 4035 TCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTA--------- 4085
Db 543 SerLeuProProArgGlyMetAlaProHisProAsnMetProGlySerGlnMetArgLeu 562
Qy 4086 ----------TCTCTCCCGTCACCGCGG 4103
Db 563 ProGlyPheAlaGlyMetIleAsnSerGluMetGluGlyProAsnValProAsnProAla 582
Qy 4104 ACAACCGCAGCAGTCATGGGATTG------CCGACCAACTCTCCTAGCATG--------- 4148
Db 583 SerArgProGlyLeuSerGlyValSerTrpProAspAspValProLysIleProAspGly 602
Qy 4149 ----------GATGGAACAGGATCATTATCTGGA---------- 4172
Db 603 ArgAsnPheProProGlyGlnGlyIlePheSerGlyProGlyArgGlyGluArgPhePro 622
Qy 4173 ----------TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGC 4211
Db 623 AsnProGlnGlyLeuSerGluGluMetPheGlnGlnGlnLeuAlaGluLysGlnLeuGly 642
Qy 4212 ACAACA------ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCG 4265
Db 643 LeuProProGlyMetAlaMetGluGlyIleArgProSerMetGluMetAsnArgMetIle 662
Qy 4266 CCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTA 4325
Db 663 ProGlySerGln---ArgHisMetGluProGlyAsnAsnProIleProProArgIlePro 681
Qy 4326 AGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAG---------- 4367
Db 682 ValGluGlyProLeuSerProSerArgGlyAspPheProLysGlyIleProProGlnMet 701


Qy 4368 ----------GAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTT 4415
Db 702 GlyProGlyArgGluLeuGlu---PheGlyMetValProSerGlyMet---------- 716
Qy 4416 AATTTTTTTAAAGACAAATCAAATATCAATTGCGTTAATAATAAGTTATATATTACATAA 4475
Db 717 ---------LysGlyAspValAsnLeuAsn---ValAsn---------- 725
Qy 4476 CTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG--- 4532
Db 726 MetGlySerAsnSerGlnMetIleProGlnLysMetArgGluAlaGlyAlaGlyProGlu 745
Qy 4532 ---------- 4532
Db 746 GluMetLeuLysLeuArgProGlyGlySerAspMetLeuProAlaGlnGlnLysMetVal 765
Qy 4533 ----------CCCATCCATTTCTTG 4547
Db 766 ProLeuProPheGlyGluHisProGlnGlnGluTyrGlyMetGlyProArgProPheLeu 785
Qy 4548 ----------AATCCAATTTCTGGAGTGATTGTTAGAGATAA---------- 4579
Db 786 ProMetSerGlnGlyProGlySerAsnSerGlyLeuArgAsnLeuArgGluProIleGly 805
Qy 4580 ----------TCTACTATTAAAATTAAACACGAAAATTCATATCCGTTAATTGAAAAT 4627
Db 806 ProAspGlnArgThrAsnSerArgLeuSerHisMetProProLeuProLeu--------- 822
Qy 4628 CACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTACAGGTGATAAC 4687
Db 823 ----------AsnProSerSerAsn 827
Qy 4688 ATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAG 4747
Db 828 ProThrSerLeuAsnThrAlaProProValGlnArg----------GlyLeuGlyArg 843
Qy 4748 GCAAATAAAGATGTACGATTTGCTGCATCC---------AGTCCTGGTTTTAACCCGCAT 4798
Db 844 LysProLeuAspIleSerValAlaGlySerGlnValHisSerProGlyIleAsnPro--- 862
Qy 4799 CCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTACCAAT 4858
Db 863 ------LeuLysSerProThrMetHisGlnValGlnSerProMetLeuGlySerProSer 880
Qy 4859 ATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTC 4918
Db 881 GlyAsnLeu---------- 883
Qy 4919 TTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGC--------- 4969
Db 884 ----------LysSerProGlnThrProSerGlnLeuAlaGlyMetLeuAla 897
Qy 4970 ---------GCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCCGCC----- 5015
Db 898 GlyProAlaAlaAlaAlaSerIleLysSerProProValLeuGly-SerAlaAlaAlaSe 917
Qy 5016 ----------ACCAAACAAGA 5026
Db 917 rProValHisLeuLysSerProSerLeuProAlaProSerProGlyTrpThrSerSerPr 937
Qy 5027 TGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTT 5086
Db 937 oLysProProLeuGlnSer----------ProGlyIleProProAsnHisLysAl 952
Qy 5087 GCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCG- 5145
Db 952 aProLeu-ThrMetAlaSerProAlaMetLeuGlyAsnValGluSerGlyGlyProProP 972
Qy 5146 ----------GGACCGTCCGACATCAATATTGAG---------- 5169
Db 972 roProThrAlaSerGlnProAlaSerValAsnIleProGlySerLeuProSerSerThrP 992
Qy 5169 ---------- 5169
Db 992 roTyrThrMetProProGluProThrLeuSerGlnAsnProLeuSerIleMetMetSerA 1012
Qy 5170 ----------CATGGAATAATTTCGGGAC 5188
Db 1012 rgMetSerLysPheAlaMetProSerSerThrProLeuTyrHisAspAlaIleLysThrV 1032
Qy 5189 TAGATGGATCAGGAATAGATACC------ATAAATCAAAATAACTGTCATTCAATGAATG 5242
Db 1032 alAlaSerSerAspAspAspSerProProAlaArgSerProAsnLeuProSerMetAsn- 1051
Qy 5243 TCGTAATGAACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTC 5302
Db 1052 --AsnMetProGlyMetGlyIleAsnThrGlnAsnProArgIle---------SerGlyP 1068
Qy 5303 CAAAT---------- 5307
Db 1068 roAsnProValValProMetProThrLeuSerProMetGlyMetThrGlnProLeuSerH 1088
Qy 5308 ----------GGACCGCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAA 5350
Db 1088 isSerAsnGlnMetProSerProAsnAlaValGlyProAsnIleProProHisGlyVal- 1107
Qy 5351 GAGAGAATTCCATAGGGTCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTG 5410
Db 1108 --------ProMetGlyProGlyLeuMetSerHisAsnProIleMetGlyHisGlySerG 1125
Qy 5411 TTCCACCTGGTGCCAGAATGATGGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATC 5470
Db 1125 lnGluProProMetValProGlnGlyArgMet------GlyPheProGlnGlyPheProP 1143
Qy 5471 CGAATATTCAGGTAAAGGCGAGTACCCCAAACACCATACAATACATGCCAGTAAGGGCAC 5530
Db 1143 ro----------ValGlnSerProProGlnGlnValP 1152
Qy 5531 AGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGG 5590
Db 1152 roPheProHisAsnGlyProSerGlyGlyGlnGlySer------PheProGlyGlyMetG 1170
Qy 5591 AATTTTTGCAGAGGTACGCTAACCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATAT 5650
Db 1170 lyPhe----------ProGlyGluGlyProLeuGlyArgProSerAsnLeu- 1183
Qy 5651 GCCCACCATCAGCCAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAG 5710
Db 1184 --ProGlnSerSerAlaAspAlaAlaLeuCysLysProGlyGlyProGlyGlyProAspS 1203
Qy 5711 CCGGAGGTATGCTAATGAATTCT---------- 5733
Db 1203 erPheThrValLeuGlyAsnSerMetProSerValPheThrAspProAspLeuGlnGluV 1223
Qy 5733 ---------- 5733
Db 1223 alIleArgProGlyAlaThrGlyIleProGluPheAspLeuSerArgIleIleProSerG 1243
Qy 5734 ----------TCCGGAGAGCAACACCAGAACAAGATCA 5761
Db 1243 luLysProSerGlnThrLeuGlnTyrPheProArgGlyGluValProGlyArgLysGlnP 1263
Qy 5762 CAAACAATCCTGGGGCAAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTA 5821
Db 1263 roGlnGlyProGlyPro------GlyPheSerHisMetGln----------GlyM 1276
Qy 5822 TTGTTGACGAAGAGGGT----------GGATTACCCGGCCATGACGGATCAATGA 5866
Db 1276 etMetGlyGluGlnAlaProArgMetGlyLeuAlaLeuProGlyMetGlyGlyProGlyP 1296
Qy 5867 ATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCA----------CATGCCA 5911
Db 1296 roValGlyThrProAspIleProLeuGlyThrAlaProSerMetProGlyHisAsnProM 1316
Qy 5912 TGCGGCCAAAT------GTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGT 5965

```
Db       1316 etArgProProAlaPheLeuGlnGlnGlyMetMetGlyProHisHisArgMetMetSerP 1336

Qy       5966 TTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACG 6025
                ||||||||:::             |||   |||:::              |
Db       1336 roAlaGlnSerThr------------MetProGlyGlnProThrLeuMetSerAsnProA 1352

Qy       6026 CTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTA 6085
                ||:::
Db       1352 laAlaAla---------------------------------------------------- 1354

Qy       6086 AGACACAACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAG 6145
                        :::      ||||||   |||      :::      :::||||||
Db       1355 -----------ValGlyMetIlePro---GlyLysAspArgGlyProAlaGlyLeuTyrT 1370

Qy       6146 TGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTG 6205
                    |||   :::         |||         :::         |||  ::::::|
Db       1370 hrHisProGlyProValGlySerProGlyMetMetMetSerMet---GlnGlyMetMetG 1389

Qy       6206 GACCTACTAATAATAATTTA-----------ATGTCAACTGCCGGAAGTGTCAGTGCTA 6253
                |||||      |||   :::           |||         |||  ::::::|||
Db       1389 lyProGlnGlnAsnIleMetIleProProGlnMetArgProArgGly---MetAlaAlaA 1408

Qy       6254 CTAACGGTGTCTCTGGCATCAATTTCGTAGGTCCCTCTTCT 6294
                  |||:::   |||   :::      ||||||   :::
Db       1408 spValGlyMetGlyGlyPheSer---GlnGlyProGlyAsn 1420

RESULT 13
AAU78460
ID   AAU78460 standard; protein; 1494 AA.
XX
AC   AAU78460;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Mouse beta-catenin nuclear localised protein.
XX
KW   Mouse; beta-catenin nuclear localised protein; cancer; gene therapy; EST;
KW   expressed sequence tag.
XX
OS   Mus musculus.
XX
PN   WO200224738-A1.
XX
PD   28-MAR-2002.
XX
PF   19-SEP-2001; 2001WO-JP008140.
XX
PR   22-SEP-2000; 2000JP-00287876.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Akiyama T,  Adachi S;
XX
DR   WPI; 2002-330014/36.
DR   N-PSDB; ABK47631.
XX
PT   New beta-catenin nuclear localized protein for diagnosis and treatment of
PT   diseases associated with nuclear localization of beta-catenin e.g.
PT   cancer.
XX
PS   Claim 1; Page 81-88; 113pp; Japanese.
XX
CC   The invention relates to a beta-catenin nuclear localised protein and DNA
CC   encoding the protein. The protein and encoding DNA are applicable in
CC   diagnosis and treatment of diseases associated with nuclear localisation
CC   of beta-catenin e.g. cancer, including gene therapy. The present sequence
CC   represents the amino acid sequence of mouse beta-catenin nuclear
CC   localised protein
XX
SQ   Sequence 1494 AA;

Alignment Scores:
Pred. No.:              1.03e-15       Length:        1494
```

```
Score:                  296.50         Matches:       308
Percent Similarity:     30.65%         Conservative:  164
Best Local Similarity:  20.00%         Mismatches:    534
Query Match:            2.46%          Indels:        535
DB:                     5              Gaps:          79

US-10-664-859-1 (1-6909) x AAU78460 (1-1494)

Qy       2484 TGCAATCAGTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTGCAATTTCGAGC 2543
                ||||||    :::|||    :::          |||              :::
Db        179 CysAsnValAlaAspProAlaMetValThrProGlnLeuGlyProGlyGlnThrAlaGln 198

Qy       2544 ATATTAGTAAGCGGAGTA--------------------GGACCAATACCCGGAATCGGA 2582
                :::   :::|||                       ||||||   ||||||:::
Db        199 LeuProLeuSerGluSerSerAlaProGlyProGlnHisGlyProGlnProGlyLeu--- 217

Qy       2583 GTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCCTCGGGTAGC 2642
                                                          :::      |||
Db        218 ----------------------------------ArgProAspValProGlyGlyGly 225

Qy       2643 AGTAATTGTTTGGATTACATGCAACAGCAAAATCACATATTCGTGTTTTCAACTCAGCTG 2702
                    :::                          ::::::||||||:::|||   |||
Db        226 GlyGlyGlyValProGlyLysProProSerGlnPheValTyrValPheThrThrHisLeu 245

Qy       2703 GCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCAC 2762
                ||||||      ||||||:::||||||   |||:::   ::::::|||:::|||||||||
Db        246 AlaAsnThrAlaAlaGluAlaValLeuGlnGlyArgAlaGluSerIleLeuAlaTyrHis 265

Qy       2763 TGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTTTTTATGAAAAACCCTTTAAAG 2822
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ------------------------------------------------------------ 265

Qy       2823 ATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATAGGCATGGGGCAGGTT 2882
                                                    :::|||         :::   |||
Db        266 ------------------------GlnGlnAsnValProArgAlaLysLeuAspGlnAla 277

Qy       2883 GGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAGCCA---------CAT 2933
                      ||||||   |||                        :::|||         :::
Db        278 ProLysValProProThrPro--------------------GluProLeuProLeuAsn 290

Qy       2934 ACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAACAGCAAACGTAGT 2993
                |||   :::   |||       :::   |||                          :::
Db        291 ThrProSerAlaGlyThrProGlnSerGlnProProProLeuProProProProProAla 310

Qy       2994 ACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGATCCTATGGGCAACGAAACT 3053
                    ||||||||      :::          :::       |||
Db        311 ProGlySerAlaProPro---AlaLeuProProGluGlyProProGluAspThrSerGln 329

Qy       3054 GAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAACTCACGA 3113
                :::|||                  |||:::::::::|||                  :::
Db        330 AspLeuAlaProAsnSerValGlyAlaAlaSerThrGlyGlyGlyThrGlyGlyThrHis 349

Qy       3114 AATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTGGAAGCA 3173
                   :::      :::   :::|||::::::
Db        350 ProAsnThrProThrAlaAlaThrAlaAsn---------------------------- 359

Qy       3174 GCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGCCTGACAACTGAAAAC 3233
                                    :::      |||   |||||||||           :::
Db        360 --------------AsnProLeuProProGlyGlyAspProGlySerAlaProGlySer 374

Qy       3234 AACATTGTATCACTGCAAGGAGTTAAGGTTCCA-------------------------- 3266
                      :::|||   |||         |||
Db        375 --------AlaLeuLeuGlyGluAlaThrProThrGlyAsnGlyGlnArgAsnLeuVal 391

Qy       3267 ---GACGAAAACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATA 3323
                   |||   |||:::   :::|||   :::|||||||||       |||      :::
Db        392 GlySerGluGlyLeuSerLysGluGlnLeuGluHisArgGluArgSerLeuGlnThrLeu 411

Qy       3324 AAAAAAATGAATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGC 3383
                :::   :::   :::   |||           |||
Db        412 ArgAspIleGluArgLeuLeuArgSerGlyGluThrGlu----------------- 425

Qy       3384 TCACAGATAACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGA 3443
```


Db 426 ProPheLeuLysGlyProProGlyGlyAlaGly--- 436
Qy 3444 TCTATTATAAATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAGCTC 3503
Db 437 GluGlyGlyProProAlaGlnAla 444
Qy 3504 TTATCGGCGACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCA 3563
Db 445 ProSerAlaAlaGlnProProProSerAlaPro---------ProGlyGlyLeu------ 459
Qy 3564 GATATGGGTGCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCT 3623
Db 460 LysLysTyrGluGluProLeuGln--------- 467
Qy 3624 GGAGTAGGTGTTGTCACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCA 3683
Db 468 ------SerMetIleSerGlnThrGlnSerLeuGlyGlyProProLeuGluHis---Glu 484
Qy 3684 AGCTCCGGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTT 3743
Db 485 ValProGlyHisProGlnGlyGlyAspMetGlyGlnGlnMetAsnMetMet--------- 501
Qy 3744 GGCAACACAAGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAAT 3803
Db 502 MetGlnArgLeuGly 506
Qy 3804 CAAGATAGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATT 3863
Db 507 GlnAspSerLeuThrProGlu-----------------GlnValAlaTrpArgLysLeu 520
Qy 3864 CAACATCAATTTTTCGAAGAACGC----------CTCAAGGGGGGC 3899
Db 521 GlnGluGluTyrTyrGluGluLysArgArgLysGluGluGlnIleGlyLeuHisGlyGly 540
Qy 3900 AAGCCCAGACAAGTCACTGGAACTGTAGTACCACAACAGCAAACCCCTTCTGGATCTGGT 3959
Db 541 ArgPro----------LeuGlnAspMetValGlyMetGly 550
Qy 3960 GGAAACTCGTTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCACTCC 4019
Db 551 Gly---------MetMetGlyArgGlyProProProProTyrHisSer 563
Qy 4020 ATCCAGAGATCTGCGTCAGTACCAATAGCCACT 4052
Db 564 LysProGlyAspGlnCysAlaProGlyMetGlyAlaGlnLeuArgGlyProMetAspVal 583
Qy 4053 CAATCGCCC---------AATCCCTCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGG 4103
Db 584 GlnAspProMetGlnLeuArgProGlyProPro----------PheProGlyProArg 599
Qy 4104 ACAACCGCAGCAGTCATGGGATTGCCGACCAACTCT----------CCTAGCATGGAT 4151
Db 600 ---------PheProGlyAsnGlnMetGlnArgValProGlyPheGly 612
Qy 4152 GGAACAGGATCATTATCTGGATCTGTTCCG---CAAGCTAATACTTCGACGGTTCAGGCA 4208
Db 613 GlyMetGlnSerMetProMetGluValProMetAsnAlaMetGlnArgProValArgPro 632
Qy 4209 GGCACAACA----------ACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCA 4253
Db 633 GlyMetAlaTrpAsnGluAspLeuProProIleGlyGlyProSerAsnPheAlaGlnAsn 652
Qy 4254 GACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTT--- 4310
Db 653 AlaValProTyrProGlyGlyGlnGlyGluAlaGluArgPheMetThrProArgValArg 672
Qy 4311 ---------CTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCC 4361
Db 673 GluGluLeuLeuArgHisGlnLeuLeuGluLysArgSerMetGlyMetGln--------- 689
Qy 4362 CCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTT 4421


Db 690 ArgProLeuGlyMetAlaGlySerGlyMet---------- 699
Qy 4422 TTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGA 4481
Db 700 GlyGlnSerMetGlu----------MetGlu 706
Qy 4482 AATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGA---------- 4529
Db 707 ArgMetIleGlnAlaHisArgGlnMetAspProAlaMetPheProGlyGlnMetThrGly 726
Qy 4529 ---------- 4529
Db 727 GlyAspGlyLeuAlaGlyThrProMetGlyIleGluPheGlyGlyGlyArgGlyLeuLeu 746
Qy 4530 ---CCGCCC----------ATCCATTTCTTGAATCCAATTTCTGGA--------- 4562
Db 747 SerProProMetGlyGlnSerGlyLeuArgGluValAspProProMetGlyProGlyAsn 766
Qy 4563 ------GTGATTGTTAGAGATAATCTACTATTAAAATTAAACACGAAAATTCATATCCGT 4616
Db 767 LeuAsnMetAsnMetAsnValAsnMetAsnMetAsnMetAsnLeuAsnValGlnMet-Th 786
Qy 4617 TAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTTCTA 4676
Db 786 rProGlnGlnGlnMetLeuMetSerGlnLysMet-Arg----------GlyP 800
Qy 4677 CAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAAATA 4736
Db 800 roGlyAspMetMet----------GlyProGlnGly---LeuSerPro------- 811
Qy 4737 GTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTTTTAACCCGC 4796
Db 812 --------GluGluMetAlaArgValArg---AlaGlnAsnSerSerGlyMetMetGlyG 828
Qy 4797 ATCCACATATG----------CAAAGCAATTCAAATT 4823
Db 828 lyProGlnLysMetLeuMetProSerGlnPheProAsnGlnGlyGlnGlnGlyPheSerG 848
Qy 4824 CAGCATTAAACGCCTATAAA----------ATGGGCTCTACCAATATACAGATGG 4868
Db 848 lyGlyGlnGlyProTyrGlnAlaMetProGlnAspMetGlyAsnThrPro---------- 864
Qy 4869 AGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAATTTATCTTCTTTTTCAGCG 4928
Db 865 ----------AspMetPheSerP 869
Qy 4929 T----CAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGCGCTCCGATAATATTC 4984
Db 869 roAspGlnSerSerValProMetGlyThrValGlyThrAlaArgLeuSer---HisMetP 888
Qy 4985 CGCTAAATCCCAATAGTGGCAATCGGCCGCCACCAAACAAGATGACCCAAAACTTCGATC 5044
Db 888 roLeuProProAlaSer---AsnProProGlySerValHisLeuAlaSerAsn------- 904
Qy 5045 CAATCTCTTCTTTGGCACAAATGTCCCAACAACTAACAAGTTGCGTGTCCAGCATGGGTA 5104
Db 905 -----ArgGlyLeuGlyArgArgProSerAspLeuThrIleSerIleAsnGlnMetGlyS 923
Qy 5105 GTCCAGCCGGAACTGGTGGTATGACGATGATGGGGGGTCCGGGACCGTCCGACATCAATA 5164
Db 923 erPro----------GlyMetGlyHisLeuLysSerProThrLeuSerGlnValHis- 938
Qy 5165 TTGAGCATGGAATAATTTCGGGACTAGATGGATCAGGAATAGATACCATAAATCAAAATA 5224
Db 939 ----------SerProLeuValThrSerProSerAlaAsnLeuLysSerProG 953
Qy 5225 ACTGTCATTCAATGAATGTCGTAATGAACTCAATGGGTCCC---------- 5265
Db 953 lnThrProSerGlnMetValProLeuProSerAlaAsnProProGlyProLeuLysSerP 973
Qy 5266 ----------CGAATGC 5272
Db 973 roGlnValLeuSerSerSerLeuGlyValArgSerProThrGlySerProSerArgLeuL 993

```
Qy       5273 TGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACCGCCTGGCTTT------AATC 5326
Db        993 ysSerProSerMetAlaValProSer--------------ProGlyTrpValAlaSerP 1008

Qy       5327 CTAATTCC------CCCAATGGTGGATTAAGAGAGAAT---------------TCCATAG 5365
Db       1008 roLysThrAlaMetProSerProGlyValSerGlnAsnLysGlnProProLeuSerIleA 1028

Qy       5366 GGTCTGGCTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCC- 5424
Db       1028 snSerSerSerThrLeuGlyAsnValGlu------GlnGlyAlaLeuProProSerAlaP 1046

Qy       5425 -----AGAATGATGGGTCGAATGCCAGTCAATTTTGGTTCGAAT---TTCAATCCGAATA 5476
Db       1046 roArgAsnSerSerSerAlaProProAlaAsnProSerSerGlyLeuMetAsnPro---- 1064

Qy       5477 TTCAGGTAAAGGCGAGTACCCCAAACACCATACAATACATGCCAGTAAGGGCACAGAACG 5536
Db       1065 -------------SerLeuProPheThrSerSerProAspProThrProSerGlnAsn- 1079

Qy       5537 CCAACAACAATAACAACAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGGAATTTT 5596
Db       1080 ----------------------------------------ProLeuSerLeuMetMetS 1086

Qy       5597 TGCAG------AGGTACGCTAACCCTCAAATGGGTGCTGTAGGCAATGGGTCG------- 5643
Db       1086 erGlnMetSerLysTyrAlaMetProSerSerThrProLeuTyrHisAsnAlaIleLysT 1106

Qy       5644 ------------------------------------CCAATATGCCCACCATCAGCCA 5665
Db       1106 hrIleAlaThrSerAspAspGluLeuLeuProAspArgProLeuLeuProProProPro- 1125

Qy       5666 GCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAA 5725
Db       1126 --------------------ProProGlnGlySerGlyProGlyIleSerAsn------- 1136

Qy       5726 TGAATTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAAT- 5784
Db       1137 --AsnGlnProAsnGlnMetHisMet--------------AsnProAlaAlaAlaGlnS 1151

Qy       5785 --------GGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGG 5836
Db       1151 erProMetGlyMetAsn---------------------------------------- 1156

Qy       5837 GTGGATTACCCGGCCATGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCA 5896
Db       1157 -----LeuProGlyGlnGlnProLeuSerHisGluProProProThrMetLeuProSerP 1175

Qy       5897 TGCGTCCACATGCC------------ATGCGGCCAAATGTAATGGGTGCGCGGATGCCAC 5944
Db       1175 roThrProLeuGlySerAsnIleProLeuHisProAsnAlaGlnGlyThrGlyGlySerS 1195

Qy       5945 CCGTTAACAGGCAAATTCAGTTTGCACAGTCATCGGATGGTATTGACTGTGTCGGGGATC 6004
Db       1195 erGlnAsnSerMetMetMet--------------------------AlaProGlyGlyP 1206

Qy       6005 CGTCATCATTTTTCACTAACGCTTCCTGCAACAGCGCTGGACCACACATGTTTGGATCAG 6064
Db       1206 roAspSerLeu------AsnAlaProCys---------GlyProValProSerSerSerG 1221

Qy       6065 CACAACAGGCCAATCAGCCTAAGACACAACAC-------------------------- 6096
Db       1221 lnMetMetSerPheProProArgLeuGlnGlnProHisGlyAlaMetAlaProThrGlyA 1241

Qy       6097 -----------------ATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGAC 6139
Db       1241 laGlyGlyProGlyLeuGlnGlnHisTyrProSerGlyMetAlaLeuProProGluAspL 1261

Qy       6140 TTGCAGTGGCACAAGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTT 6199
Db       1261 euPro---ThrGlnProProGlyProIleProProGlnGlnHisLeuMetGlyLysGlyM 1280

Qy       6200 TAATTGGACCTACTAATAATAATTTAATGTCAACTGCCGGAAGTGTCAGTGCTACTAACG 6259
Db       1280 etThrGly--------------ArgMetGlyAspAlaTyrProProGlyValLeuProG 1295

Qy       6260 GTGTCTCTGGCATC--------------------AATTTCGTAGGTCCCTCTTCTACG- 6297
Db       1295 lyValAlaSerValLeuAsnAspProGluLeuSerGluValIleArgProThrProThrG 1315

Qy       6298 --------------GACCTGAAGTAT---------GCCCAGCAATATCATAGTTTTCAGC 6334
Db       1315 lyIleProGluPheAspLeuSerArgIleIleProSerGluLysProSerSerThrLeuG 1335

Qy       6335 AGCAGTTATATGCTACCAACACCAGAAGTCAA------------CAACAACAGCATATGC 6382
Db       1335 ln------TyrPheProLysSerGluAsnGlnProProLysAlaGlnProProAsnLeuH 1353

Qy       6383 ACCAGCAGCACCAGAGCAACATGATA--------------ACAATGCCGCCGAATTTAT 6427
Db       1353 isLeuMetAsnLeuGlnAsnMetMetAlaGluGlnThrProSerArgProProAsnLeu- 1372

Qy       6428 CACCAAATCCAACGTTCTTTGTCAACAAATAAACTTCTAAATTTTTGCCGCCCT 6481
Db       1373 --ProGly-GlnGlnGlyValGlnArgGlyLeuSerMetSerMetCysHisPro 1389
```

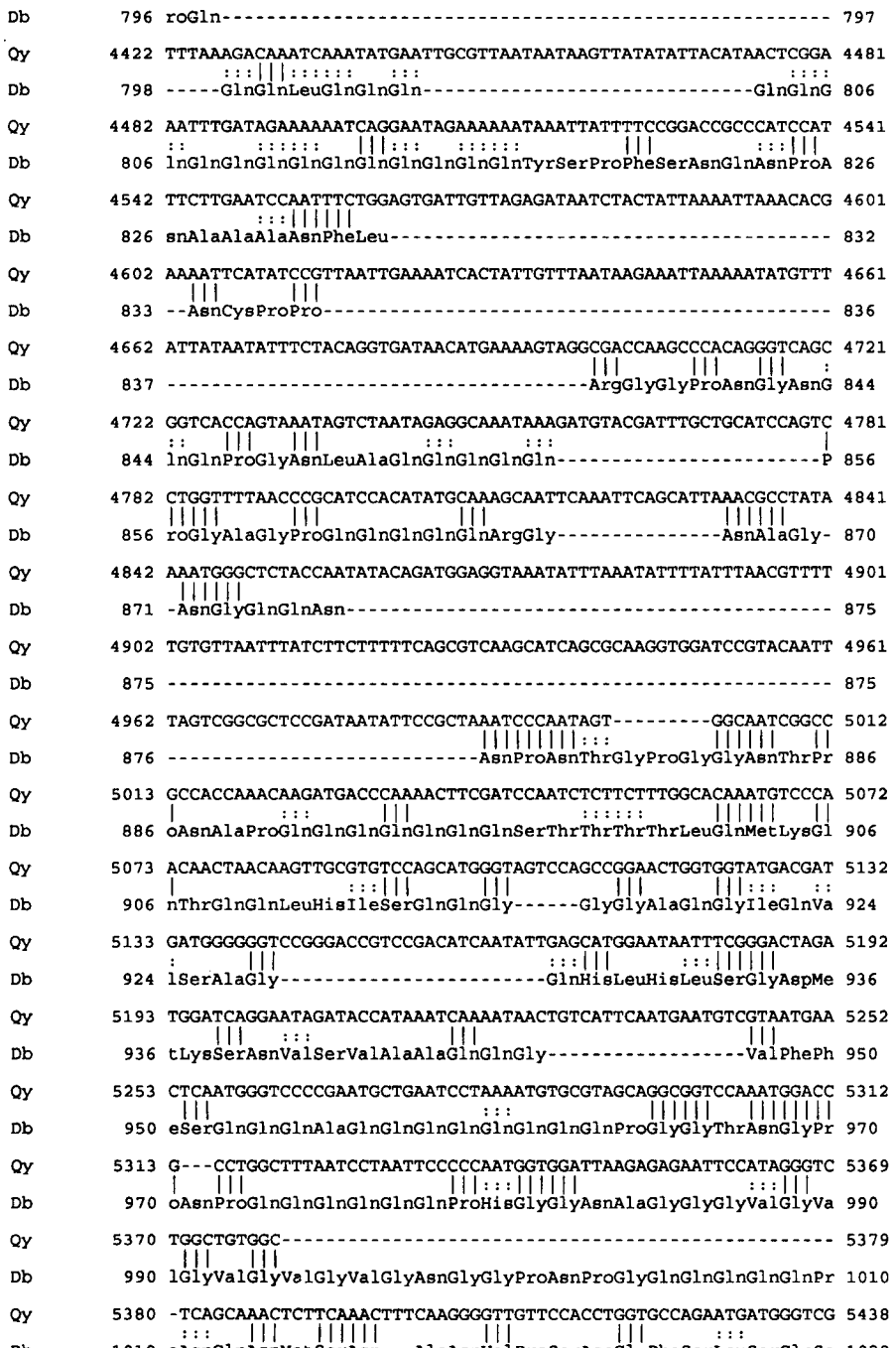
RESULT 14
AAE36095
ID AAE36095 standard; protein; 1596 AA.
XX
AC AAE36095;
XX
DT 26-JUN-2003 (first entry)
XX
DE Drosophila sp. Mastermind (Mam) protein.
XX
KW Notch signal transduction; Mastermind; Mam; Mip1; Mip30; Mip6; cancer;
KW gene therapy; fruitfly; cytostatic.
XX
OS Drosophila sp.
XX
PN WO2002102987-A2.
XX
PD 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WO-US019189.
XX
PR 18-JUN-2001; 2001US-0299112P.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Artavanis-Tsakonas S, Lake RJ;
XX
DR WPI; 2003-167509/16.
DR N-PSDB; AAD54620.
XX
PT Inhibiting or agonizing Notch signal transduction in a cell for preparing
PT a composition for treating or preventing cancer comprises contacting the
PT cell with an antagonist or agonist of sumolation.
XX
PS Disclosure; Fig 1; 84pp; English.
XX
CC The invention relates to a method for modulating Notch signal
CC transduction in a cell which comprises contacting the cell with an
CC antagonist or agonist of sumolation. The invention is based on the
CC interactions of Mastermind (Mam) protein with the Mip1, Mip30 and Mip6
CC proteins. The method is useful for preparing a composition for treating
CC or preventing a disease e.g. cancer. The invention is also used in gene
CC therapy. The present sequence is Drosophila sp. Mam protein used to
CC illustrate the method of the invention
XX
SQ Sequence 1596 AA;

Alignment Scores:
Pred. No.: 1.84e-13 Length: 1596

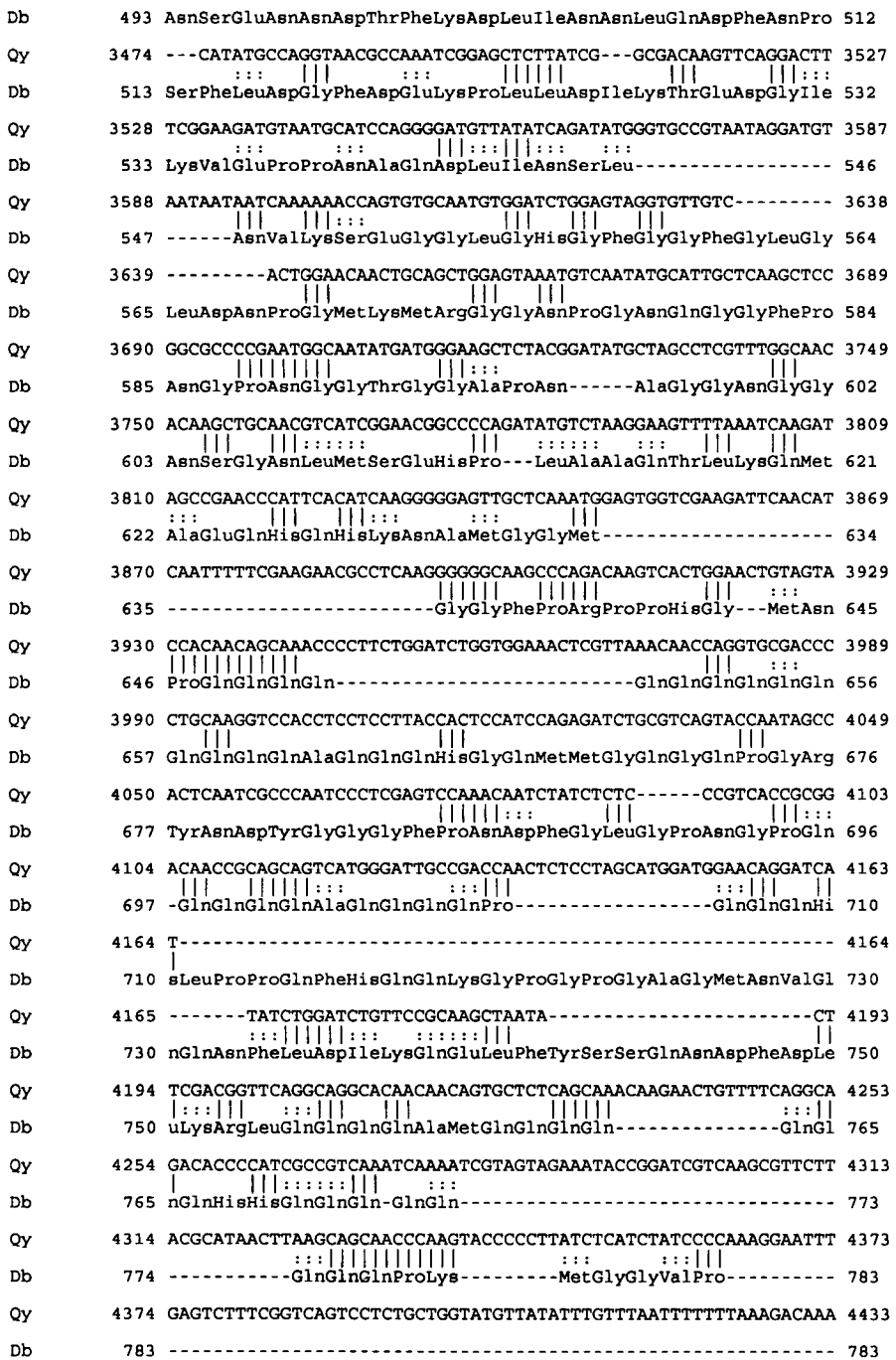
Score:	270.50	Matches:	299
Percent Similarity:	29.56%	Conservative:	148
Best Local Similarity:	19.78%	Mismatches:	544
Query Match:	2.25%	Indels:	521
DB:	6	Gaps:	70

US-10-664-859-1 (1-6909) x AAE36095 (1-1596)

```
Qy       2658 TACATGCAACAGCAAAATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCC 2717
Db        203 HisGlnGlnGlnGlnGlnHis-----------GlnGlnHisGlnGlnHisGlnGlnAla 218

Qy       2718 GAATCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCT 2777
Db        219 GlnThrMetLeuAlaGlyGlnLeuGlnSerSerVal------HisValGlnGlnLysPhe 236

Qy       2778 ACAAAAAGCTTCCTGGAAGACTTTTTT-------------------------ATGAAA 2810
Db        237 LeuLysArgProAlaGluAspValAspAsnGlyProAspSerPheGluProProHisLys 256

Qy       2811 AACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATAGGC 2870
Db        257 LeuProAsnAsnAsnAsnAsnSerAsnSerAsnAsnAsnGlyAsnAlaAsnAlaAsn 276

Qy       2871 ATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTA---------------------- 2906
Db        277 AsnGlyGlyAsnGlySerAsnThrGlyAsnAsnThrAsnAsnAsnGlyAsnSerThrAsn 296

Qy       2907 ----------------------------------------------------GCCAAAATA 2915
Db        297 AsnAsnGlyGlySerAsnAsnAsnGlySerGluAsnLeuThrLysPheSerValGluIle 316

Qy       2916 ACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACAT 2975
Db        317 ValGlnGlnLeuGluPheThrThrSerProAlaAsnSerGlnProGln------------ 332

Qy       2976 GAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGAT 3035
Db        333 GlnIleSerThrAsnValThrValLysAlaLeuThrAsnThrSerVal---LysSerGlu 351

Qy       3036 CCT-------------------------------------ATGGGCAACGAAACTGAA 3056
Db        352 ProGlyValGlyGlyGlyGlyGlyGlyGlyGlyGlyAsnSerGlyAsnAsnAsn--- 370

Qy       3057 TTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAACTCACGAAAT 3116
Db        371 ------------AsnAsnGlyGlyGlyGlyGlyGlyGlyAsnGlyAsnAsnAsnAsnAsn 386

Qy       3117 CATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTGGAAGCAGCT 3176
Db        387 GlyGlyAspHisHisGlnGlnGlnGlnGlnHisGlnHisGlnGlnGlnGlnGlnGln 406

Qy       3177 GGCGTTGATTTGGGACAGGTC---------ACAAAAGGAAGCGATCCTGGC------CTG 3221
Db        407 GlyGlyGlyLeuGlyGlyLeuGlyAsnAsnGlyArgGlyGlyGlyProGlyGlyMetAla 426

Qy       3222 ACAACTGAAAACAACATTGTA---TCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTT 3278
Db        427 ThrGlyProGlyGlyValAlaGlyGlyLeuGlyGlyMetGlyMetProProAsnMetMet 446

Qy       3279 ACACCACAACAGCGGCAA------CATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATG 3332
Db        447 SerAlaGlnGlnLysSerAlaLeuGlyAsnLeuAlaAsnLeuValGluCysLysArgGlu 466

Qy       3333 AATCAATTTCTTTTTCCTGAA-----------AATGAGAATTCAGTAGGAGCTAATGTA 3380
Db        467 ProAspHisAspPheProAspLeuGlySerLeuAspLysAspGlyGlyGlyGly------ 484

Qy       3381 AGCTCACAGATAACAAAAATTCCAGGA------GATTTAATGATGGGGATGTCGGGTGGC 3434
Db        485 --------------GlnPheProGlyPheProAspLeuLeu--------------Gly 494

Qy       3435 GGAGGCGGATCTATTATAAATCCGACGATGCGACAACTG---------------------- 3473

Db        495 AspAspAsnSerGluAsnAsnAspThrPheLysAspLeuIleAsnAsnLeuGlnAspPhe 514

Qy       3474 ---------CATATGCCAGGTAACGCCAAATCGGAGCTCTTATCG---GCGACAAGTTCA 3521
Db        515 AsnProSerPheLeuAspGlyPheAspGluLysProLeuLeuAspIleLysThrGluAsp 534

Qy       3522 GGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATA 3581
Db        535 GlyIleLysValGluProProAsnAlaGlnAspLeuIleAsnSerLeu------------ 550

Qy       3582 GGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTC--- 3638
Db        551 ------------AsnValLysSerGluGlyGlyLeuGlyHisGlyPheGlyGlyPheGly 566

Qy       3639 ---------------ACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCA 3683
Db        567 LeuGlyLeuAspAsnProGlyMetLysMetArgGlyGlyAsnProGlyAsnGlnGlyGly 586

Qy       3684 AGCTCCGGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTT 3743
Db        587 PheProAsnGlyProAsnGlyGlyThrGlyGlyAlaProAsn------AlaGlyGlyAsn 604

Qy       3744 GGCAACACAAGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAAT 3803
Db        605 GlyGlyAsnSerGlyAsnLeuMetSerGluHisPro---LeuAlaAlaGlnThrLeuLys 623

Qy       3804 CAAGATAGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATT 3863
Db        624 GlnMetAlaGluGlnHisGlnHisLysAsnAlaMetGlyGlyMet-------------- 638

Qy       3864 CAACATCAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACT 3923
Db        639 ----------------------------GlyGlyProArgProProHisGly--- 647

Qy       3924 GTAGTACCACAACAGCAAACCCCTTCT---------------------------------- 3950
Db        648 MetAsnProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnAlaGlnGln 667

Qy       3951 ------------------GGATCTGGT-----------------------GGAAACTCG 3968
Db        668 GlnHisGlyGlnMetMetGlyGlnGlyGlnProGlyArgTyrAsnAspTyrGlyGlyGly 687

Qy       3969 TTAAACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGA 4028
Db        688 PheProAsnAspPheGly------LeuGlyProAsnGlyProGlnGlnGln---GlnGln 704

Qy       4029 TCTGCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCT 4088
Db        705 GlnAla-GlnGlnGlnGlnProGlnGlnGlnHisLeuPro------------------- 717

Qy       4089 CTCCCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATG 4148
Db        718 ------------------ProGlnPheHisGlnGlnLysGlyProGlyProGlyAlaGl 731

Qy       4149 GATGGAA---CAGGATCATTATCTGGATCTGTTCCGCAAGCTAATA------------- 4191
Db        731 yMetAsnValGlnGlnAsnPheLeuAspIleLysGlnGluLeuPheTyrSerSerGlnAs 751

Qy       4192 ----------CTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCAAACAAGAAC 4241
Db        751 nAspPheAspLeuLysArgLeuGlnGlnGlnGlnAlaMetGlnGlnGlnGln-------- 768

Qy       4242 TGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCG 4301
Db        769 -------GlnGlnGlnHisHisGlnGlnGlnGlnProLysMetGlyGlyValProAsn-- 785

Qy       4302 TCAAGCGTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCC 4361
Db        786 ---------------PheAsn-LysGlnGlnGlnGln-----------GlnGlnValP 796

Qy       4362 CCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTT 4421
```


Db 796 roGln 797
Qy 4422 TTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGA 4481
Db 798 -----GlnGlnLeuGlnGlnGln-----GlnGlnG 806
Qy 4482 AATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCAT 4541
Db 806 lnGlnGlnGlnGlnGlnGlnGlnGlnGlnTyrSerProPheSerAsnGlnAsnProA 826
Qy 4542 TTCTTGAATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAATTAAACACG 4601
Db 826 snAlaAlaAlaAsnPheLeu 832
Qy 4602 AAAATTCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTT 4661
Db 833 --AsnCysProPro 836
Qy 4662 ATTATAATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGC 4721
Db 837 ArgGlyGlyProAsnGlyAsnG 844
Qy 4722 GGTCACCAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTC 4781
Db 844 lnGlnProGlyAsnLeuAlaGlnGlnGlnGlnGln-----P 856
Qy 4782 CTGGTTTTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATA 4841
Db 856 roGlyAlaGlyProGlnGlnGlnGlnArgGly-----AsnAlaGly- 870
Qy 4842 AAATGGGCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTT 4901
Db 871 -AsnGlyGlnGlnAsn 875
Qy 4902 TGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATT 4961
Db 875 875
Qy 4962 TAGTCGGCGCTCCGATAATATTCCGCTAAATCCCAATAGT---------GGCAATCGGCC 5012
Db 876 -----AsnProAsnThrGlyProGlyGlyAsnThrPr 886
Qy 5013 GCCACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCA 5072
Db 886 oAsnAlaProGlnGlnGlnGlnGlnGlnGlnSerThrThrThrThrLeuGlnMetLysGl 906
Qy 5073 ACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGAT 5132
Db 906 nThrGlnGlnLeuHisIleSerGlnGlnGly------GlyGlyAlaGlnGlyIleGlnVa 924
Qy 5133 GATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGA 5192
Db 924 lSerAlaGly-----GlnHisLeuHisLeuSerGlyAspMe 936
Qy 5193 TGGATCAGGAATAGATACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAA 5252
Db 936 tLysSerAsnValSerValAlaAlaGlnGlnGly-----ValPhePh 950
Qy 5253 CTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACC 5312
Db 950 eSerGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnProGlyGlyThrAsnGlyPr 970
Qy 5313 G---CCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTC 5369
Db 970 oAsnProGlnGlnGlnGlnGlnProHisGlyGlyAsnAlaGlyGlyGlyValGlyVa 990
Qy 5370 TGGCTGTGGC 5379
Db 990 lGlyValGlyValGlyValGlyAsnGlyGlyProAsnProGlyGlnGlnGlnGlnPr 1010
Qy 5380 -TCAGCAAACTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGGTCG 5438
Db 1010 oAsnGlnAsnMetSerAsn---AlaAsnValProSerAspGlyPheSerLeuSerGlnSe 1029
Qy 5439 AATGCCAGTCAATTTTGGT-----TCGAATTT 5465
Db 1029 rGlnSerMetAsnPheAsnGlnGlnGlnGlnGlnAlaAlaAlaGlnGlnGlnGlnVa 1049
Qy 5466 CAATCCGAATATT---------CAGGTAAAGGCGAGTACCCCAAACACCATACAATACAT 5516
Db 1049 lGlnProAsnMetArgGlnArgGlnThrGlnAlaGlnAlaAlaAlaAlaAlaAlaAlaAl 1069
Qy 5517 GCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAAT 5576
Db 1069 aAlaAlaGlnAlaGlnAlaAlaAlaAsnAlaSerGly 1081
Qy 5577 GCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAACCCTCAAATGGGT----------- 5625
Db 1082 ----ProAsnValProLeuMetGlnGln---------ProGlnValGlyValGlyValGl 1097
Qy 5626 ----------GCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGC 5663
Db 1097 yValGlyValGlyValGlyValGlyValGlyAsnGlyGlyValValGlyGlyProGlySe 1117
Qy 5664 CAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCA------GGAGCCGGAGG 5717
Db 1117 rGlyGlyProAsnAsnGlyAlaMetAsnGlnMetGlyGlyProMetGlyGlyMetProGl 1137
Qy 5718 TATGCTAATGAATTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGC 5777
Db 1137 yMetGlnMetGlyGlyPro-----MetAsnProMetGlnMetAsnProAsnAl 1153
Qy 5778 AAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGG 5837
Db 1153 aAlaGly-----ProThrAlaGlnGlnMetMetMetGlySerGlyAlaGl 1168
Qy 5838 TGGA---------TTACCCGGCCATGACGGATCAATGAAT----------AT 5870
Db 1168 yGlyProGlyGlnValProGlyProGlyGlnGlyProAsnProAsnGlnAlaLysPheLe 1188
Qy 5871 TGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCATGCGGCCAAATGTAATGGG 5930
Db 1188 uGlnGlnGlnGlnMetMetArgAlaGlnAlaMetGlnGlnGlnGlnGlnHisMetSerGl 1208
Qy 5931 TGCGCGGATGCCACCC----------GTTAACAGGCAAATTCAGTTTGCACAGTC 5975
Db 1208 yAlaArgProProProProGluTyrAsnAlaThrLysAlaGlnLeuMetGlnAlaGlnMe 1228
Qy 5976 ATCGGAT----------GG 5984
Db 1228 tMetGlnGlnThrValGlyGlyGlyGlyValGlyValGlyGlyValGlyValGlyValGl 1248
Qy 5985 TATTGACTGTGTCGGGGATCCGTCA---TCATTTTTCACTAACGCTTCCTGCAACAGCGC 6041
Db 1248 yValGlyGlyValGlyGlyAlaAsnGlyGlyArgProProAsnSerAlaAlaGlnAlaAl 1268
Qy 6042 T----------GGACCACACATG-------- 6054
Db 1268 aAlaMetArgArgMetThrGlnGlnProIleProProSerGlyProMetMetArgProGl 1288
Qy 6054 ---------- 6054
Db 1288 nHisAlaMetTyrMetGlnGlnHisGlyGlyAlaGlyGlyGlyProArgThrGlyMetGl 1308
Qy 6055 ----------TTTGGATCAGCACAACAGGCCAATCAGCC 6083
Db 1308 yValProTyrGlyGlyGlyArgGlyGlyProMetGlyGlyProGlnGlnGlnGlnArgPr 1328
Qy 6084 TAAGACACAACACATA------AAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGG 6137
Db 1328 oProAsnValGlnValThrProAspGlyMetProMetGlySerGlnGlnGluTrpArgHi 1348
Qy 6138 ACTTGCAGTGGCACAAGGGCAGATCCAACTG----------CATGGGCAAGGACATGC 6185
Db 1348 sMetMetMetThrGlnGlnGlnThrGlnMetGlyPheGlyGlyProGlyProGlyGlyPr 1368

```
Qy       6186 GCAGGGTCAGTCTTTAATTGGACCTACTAATAATAATTTAATGTCAACTGCCGGAAGTGT 6245
                 |||         |||           |||   |||
Db       1368 oMetArgGlnGlyProGlyGlyPheAsnGlyGlyAsnPheMetProAsn---------- 1384

Qy       6246 CAGTGCTACTAACGGTGTCTCTGGCATCAAT----------------TTCGTAGGTCC 6287
                  |||   ||||||   :::|||                                |||||
Db       1385 -GlyAlaProAsnGlyAlaAlaGlySerGlyProAsnAlaGlyGlyMetMetThrGlyPr 1404

Qy       6288 CTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCAGCAGCAGTTATATGC 6347
                 |:::      ::::::                  ||||||||||||
Db       1404 oAsnValProGlnMetGlnLeuThrProAla------GlnMetGlnGlnGlnLeu----- 1420

Qy       6348 TACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCACCAGAGCAACATGAT 6407
                     |||   ||||||||||||         ||||||   |||
Db       1421 -------MetArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnHis-------- 1435

Qy       6408 AACAATGCCGCCGAATTTATCACCAAAT 6435
                  |||   |||      :::   |||
Db       1436 ----MetGlyProGlyAlaAlaAsnAsn 1443

RESULT 15
ABB63688
ID   ABB63688 standard; protein; 1594 AA.
XX
AC   ABB63688;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 17856.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL07791.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 17856; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1594 AA;

Alignment Scores:
Pred. No.:              2.24e-13        Length:         1594
Score:                  269.50          Matches:        295
Percent Similarity:     29.36%          Conservative:   146
Best Local Similarity:  19.64%          Mismatches:     548
Query Match:            2.24%           Indels:         514
DB:                     4               Gaps:           67

US-10-664-859-1 (1-6909) x ABB63688 (1-1594)

Qy       2658 TACATGCAACAGCAAAATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCC 2717
              :::   |||||||||   |||                                ::::::   |||
Db        203 HisGlnGlnGlnGlnGlnHis-----------GlnGlnHisGlnGlnHisGlnGlnAla 218

Qy       2718 GAATCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCT 2777
              :::::::::|||:::||||||   |||:::   :::      |||      |||
Db        219 GlnThrMetLeuAlaGlyGlnLeuGlnSerServal------HisValGlnGlnLysPhe 236

Qy       2778 ACAAAAAGCTTCCTGGAAGACTTTTTT-------------------------ATGAAA 2810
              |||          ||||||                                   |||
Db        237 LeuLysArgProAlaGluAspValAspAsnGlyProAspSerPheGluProProHisLys 256

Qy       2811 AACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATAGGC 2870
              |||         |||           :::|||:::   |||
Db        257 LeuProAsnAsnAsnAsnAsnSerAsnSerAsnAsnAsnAsnGlyAsnAlaAsnAlaAsn 276

Qy       2871 ATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTA------------------------ 2906
              |||      |||          |||
Db        277 AsnGlyGlyAsnGlySerAsnThrGlyAsnAsnThrAsnAsnAsnGlyAsnSerThrAsn 296

Qy       2907 ----------------------------------------------------GCCAAAATA 2915
                                                                    :::|||
Db        297 AsnAsnGlyGlySerAsnAsnAsnGlySerGluAsnLeuThrLysPheSerValGluIle 316

Qy       2916 ACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAACAT 2975
              ||||||         |||   :::            :::||||||
Db        317 ValGlnGlnLeuGluPheThrThrSerAlaAlaAsnSerGlnProGln------------ 332

Qy       2976 GAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCTGAT 3035
              :::   |||         ||||||   |||   :::|||:::   |||   :::|||:::
Db        333 GlnIleSerThrAsnValThrValLysAlaLeuThrAsnThrSerVal---LysSerGlu 351

Qy       3036 CCT-------------------------------------ATGGGCAACGAAACTGAA 3056
              |||                                     ||||||
Db        352 ProGlyValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlySerGlyAsnAsnAsn--- 370

Qy       3057 TTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGTAGGTCTGGACAAAACTCACGAAAT 3116
                        |||                 :::|||   |||:::   |||
Db        371 ------------AsnAsnGlyGlyGlyGlyGlyGlyAsnGlyAsnAsnAsnAsnAsn 386

Qy       3117 CATGTAGACAGTATCAGTACATCCAGCGAGTCACAGGCAATAAAGATACTGGAAGCAGCT 3176
                  |||                 :::   |||      :::      :::
Db        387 GlyGlyAspHisHisGlnGlnGlnHisGlnGlnGlnGlnGlnGln------GlnGlyGly 404

Qy       3177 GGCGTT---GATTTGGGACAGGTCACAAAAGGAAGCGATCCTGGC------CTGACAACT 3227
              |||:::      ||||||          :::|||      ||||||          |||
Db        405 GlyLeuGlyGlyLeuGlyAsnAsnGlyArgGlyGlyGlyProGlyGlyMetAlaThrGly 424

Qy       3228 GAAAACAACATTGTA---TCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCA 3284
                    :::          |||  |||:::   :::|||          ::::::
Db        425 ProGlyGlyValAlaGlyGlyLeuGlyGlyMetGlyMetProProAsnMetMetSerAla 444

Qy       3285 CAACAGCGGCAA------CATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAA 3338
              ||||||:::                          |||   :::   |||:::
Db        445 GlnGlnLysSerAlaLeuGlyAsnLeuAlaAsnLeuValGluCysLysArgGluProAsp 464

Qy       3339 TTTCTTTTTCCTGAA------------AATGAGAATTCAGTAGGAGCTAATGTAAGCTCA 3386
                    ||||||:::           :::::::::      |||
Db        465 HisAspPheProAspLeuGlySerLeuAspLysAspGlyGlyGlyGly------------ 480

Qy       3387 CAGATAACAAAAATTCCAGGA------GATTTAATGATGGGGATGTCGGGTGGCGGAGGC 3440
                    :::   ||||||      ||||||:::                  |||
Db        481 ---------GlnPheProGlyPheProAspLeuLeu--------------GlyAspAsp 492

Qy       3441 GGATCTATTATAAATCCGACGATGCGACAACTG---------------------------- 3473
              |||      |||   |||   :::   |||
```


Db 493 AsnSerGluAsnAsnAspThrPheLysAspLeuIleAsnAsnLeuGlnAspPheAsnPro 512
Qy 3474 ---CATATGCCAGGTAACGCCAAATCGGAGCTCTTATCG---GCGACAAGTTCAGGACTT 3527
Db 513 SerPheLeuAspGlyPheAspGluLysProLeuLeuAspIleLysThrGluAspGlyIle 532
Qy 3528 TCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGT 3587
Db 533 LysValGluProProAsnAlaGlnAspLeuIleAsnSerLeu---------- 546
Qy 3588 AATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTC--------- 3638
Db 547 ------AsnValLysSerGluGlyGlyLeuGlyHisGlyPheGlyGlyPheGlyLeuGly 564
Qy 3639 ---------ACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCC 3689
Db 565 LeuAspAsnProGlyMetLysMetArgGlyGlyAsnProGlyAsnGlnGlyGlyPhePro 584
Qy 3690 GGCGCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAAC 3749
Db 585 AsnGlyProAsnGlyGlyThrGlyGlyAlaProAsn------AlaGlyGlyAsnGlyGly 602
Qy 3750 ACAAGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGAT 3809
Db 603 AsnSerGlyAsnLeuMetSerGluHisPro---LeuAlaAlaGlnThrLeuLysGlnMet 621
Qy 3810 AGCCGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACAT 3869
Db 622 AlaGluGlnHisGlnHisLysAsnAlaMetGlyGlyMet---------- 634
Qy 3870 CAATTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTA 3929
Db 635 ----------GlyGlyPheProArgProProHisGly---MetAsn 645
Qy 3930 CCACAACAGCAAACCCCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCC 3989
Db 646 ProGlnGlnGlnGln----------GlnGlnGlnGlnGlnGln 656
Qy 3990 CTGCAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCC 4049
Db 657 GlnGlnGlnGlnAlaGlnGlnGlnHisGlyGlnMetMetGlyGlnGlyGlnProGlyArg 676
Qy 4050 ACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTC------CCGTCACCGCGG 4103
Db 677 TyrAsnAspTyrGlyGlyGlyPheProAsnAspPheGlyLeuGlyProAsnGlyProGln 696
Qy 4104 ACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATGGATGGAACAGGATCA 4163
Db 697 -GlnGlnGlnGlnAlaGlnGlnGlnGlnPro----------GlnGlnGlnHi 710
Qy 4164 T---------- 4164
Db 710 sLeuProProGlnPheHisGlnGlnLysGlyProGlyProGlyAlaGlyMetAsnValGl 730
Qy 4165 -------TATCTGGATCTGTTCCGCAAGCTAATA----------CT 4193
Db 730 nGlnAsnPheLeuAspIleLysGlnGluLeuPheTyrSerSerGlnAsnAspPheAspLe 750
Qy 4194 TCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCA 4253
Db 750 uLysArgLeuGlnGlnGlnGlnAlaMetGlnGlnGlnGln----------GlnGl 765
Qy 4254 GACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTT 4313
Db 765 nGlnHisHisGlnGlnGln-GlnGln---------- 773
Qy 4314 ACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCTCATCTATCCCCAAAGGAATTT 4373
Db 774 ----------GlnGlnGlnProLys---------MetGlyGlyValPro---------- 783
Qy 4374 GAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTTTTAAAGACAAA 4433
Db 783 ---------- 783
Qy 4434 TCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGAAATTTGATAGAA 4493
Db 784 ----------AsnPheAsnLysG 788
Qy 4494 AAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCATTTCTTGAATCCA 4553
Db 788 lnGlnGlnGln---------- 791
Qy 4554 ATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAATTAAACACGAAAATTCATATC 4613
Db 791 ---------- 791
Qy 4614 CGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAAAATATGTTTATTATAATATTT 4673
Db 791 ---------- 791
Qy 4674 CTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCACAGGGTCAGCGGTCACCAGTAA 4733
Db 792 ----------GlnGlnValProGlnGlnGlnLeuGlnGlnGlnG 803
Qy 4734 ATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTGCATCCAGTCCTGGTTTTAACC 4793
Db 803 lnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnTyrSerProPheSerAsnG 823
Qy 4794 CGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAAACGCCTATAAAATGGGCTCTA 4853
Db 823 lnAsnProAsnAlaAlaAlaAsn---------PheLeuAsnCysProProArgGlyGlyP 840
Qy 4854 CCAATAT----ACAGATGGAGGTAAATATTTAAATATTTTATTTAACGTTTTTGTGTTAA 4909
Db 840 roAsnGlyAsnGlnGlnProGlyAsn---------- 848
Qy 4910 TTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATCCGTACAATTTAGTCGGC 4969
Db 849 --------LeuAlaGlnGlnGlnGlnGlnProGlyAlaGlyProGlnGlnGlnGlnGlnA 866
Qy 4970 GC---------TCCGATAATATTCCGCTAAATCCCAATAGT---------GGCAATCGGC 5011
Db 866 rgGlyAsnAlaAlaAsnGlyGlnGlnAsnAsnProAsnAlaGlyProGlyGlyAsnThrP 886
Qy 5012 CGCCACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCC 5071
Db 886 roAsnAlaProGlnGlnGlnGlnGlnSerThrThrThrThrLeu---GlnMetLysG 905
Qy 5072 AACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGA 5131
Db 905 lnThrGlnGlnLeuHisIleSerGlnGlnGly------GlyGlyAlaHisGlyIleGlnV 923
Qy 5132 TGATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAG 5191
Db 923 alSerAlaGly----------GlnHisLeuHisLeuSerGlyAspM 935
Qy 5192 ATGGATCAGGAATAGATACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGA 5251
Db 935 etLysSerAsnValSerValAlaAlaGlnGlnGly----------ValPheP 949
Qy 5252 ACTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGAC 5311
Db 949 heSerGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnProGlyGlyThrAsnGlyP 969
Qy 5312 CG---CCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGT 5368
Db 969 roAsnProGlnGlnGlnGlnGlnProHisGlyGlyAsnAlaGlyGlyGlyValGlyV 989
Qy 5369 CTGGCTGTGGC---------- 5379
Db 989 alGlyValGlyValGlyValGlyAsnGlyGlyProAsnProGlyGlnGlnGlnGlnP 1009
Qy 5380 --TCAGCAAACTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGATGGGTC 5437
Db 1009 roAsnGlnAsnMetSerAsn---AlaAsnValProSerAspGlyPheSerLeuSerGlnS 1028

```
Qy       5438 GAATGCCAGTCAATTTTGGT--------------------------------TCGAATT 5464
                  :::||||||
Db       1028 erGlnSerMetAsnPheAsnGlnGlnGlnGlnGlnAlaAlaAlaGlnGlnGlnV 1048

Qy       5465 TCAATCCGAATATT---------CAGGTAAAGGCGAGTACCCCAAACACCATACAATACA 5515
                ||||||:::         |||   :::|||
Db       1048 alGlnProAsnMetArgGlnArgGlnThrGlnAlaGlnAlaAlaAlaAlaAlaAlaA 1068

Qy       5516 TGCCAGTAAGGGCACAGAACGCCAACAACAATAACAACAATGGAGCTAATAATGTGCGAA 5575
                  :::||||||   |||   |||   :::
Db       1068 laAlaAlaGlnAlaGlnAlaAlaAlaAsnAlaSerGly---------------------- 1080

Qy       5576 TGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAACCCTCAAATGGGT---------- 5625
                  |||::::::      :::|||:::         ||||||:::|||
Db       1081 -----ProAsnValProLeuMetGlnGln---------ProGlnValGlyValGlyValG 1096

Qy       5626 ----------------------GCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAG 5662
                                    ||||||||||||      :::      |||   :
Db       1096 lyValGlyValGlyValGlyValGlyValGlyAsnGlyGlyValValGlyGlyProGlyS 1116

Qy       5663 CCAGCGACGGTACTCCTGGAATGCCAGGATTGATGGCGGGACCA------GGAGCCGGAG 5716
                ::          |||         |||   ||||||      |||      |
Db       1116 erGlyGlyProAsnAsnGlyAlaMetAsnGlnMetGlyGlyProMetGlyGlyMetProG 1136

Qy       5717 GTATGCTAATGAATTCTTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGG 5776
                |||||   |||                        |||   :::      ||||||   |
Db       1136 lyMetGlnMetGlyGlyPro------------MetAsnProMetGlnMetAsnProAsnA 1152

Qy       5777 CAAGCAATGGTATTAACTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGG 5836
                ||:::                          ||||||  :::                |
Db       1152 laAlaGly---------------ProThrAlaGlnGlnMetMetMetGlySerGlyAlaG 1167

Qy       5837 GTGGA---------TTACCCGGCCATGACGGATCAATGAAT-----------------A 5869
                |||||         :::||||||          |||                   :
Db       1167 lyGlyProGlyGlnValProGlyProGlyGlnGlyProAsnProAsnGlnAlaLysPheL 1187

Qy       5870 TTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCATGCGGCCAAATGTAATGG 5929
                ::   |||      |||:::|||                    :::   ::::::   |
Db       1187 euGlnGlnGlnGlnMetMetArgAlaGlnAlaMetGlnGlnGlnGlnHisMetSerG 1207

Qy       5930 GTGCGCGGATGCCACCC---------------GTTAACAGGCAAATTCAGTTTGCACAGT 5974
                ||||||||   ||||||                    |||:::      ||||||
Db       1207 lyAlaArgProProProProGluTyrAsnAlaThrLysAlaGlnLeuMetGlnAlaGlnM 1227

Qy       5975 CATCGGAT-----------------------------------------------------G 5983
                                                                          |
Db       1227 etMetGlnGlnThrValGlyGlyGlyGlyValGlyValGlyGlyValGlyValGlyValG 1247

Qy       5984 GTATTGACTGTGTCGGGGATCCGTCA---TCATTTTTCACTAACGCTTCCTGCAACAGCG 6040
                ||:::      ||||||      :::         |||   |||::::::      :::|
Db       1247 lyValGlyGlyValGlyGlyAlaAsnGlyGlyArgPheProAsnSerAlaAlaGlnAlaA 1267

Qy       6041 CT-------------------------------------GGACCACACATG------- 6054
                ||                                   ||||||   |||
Db       1267 laAlaMetArgArgMetThrGlnGlnProIleProProSerGlyProMetMetArgProG 1287

Qy       6054 ------------------------------------------------------------ 6054

Db       1287 lnHisAlaMetTyrMetGlnGlnHisGlyGlyAlaGlyGlyGlyProArgThrGlyMetG 1307

Qy       6055 -------------------------------TTTGGATCAGCACAACAGGCCAATCAGC 6082
                                              |||      ||||||      :::|
Db       1307 lyValProTyrGlyGlyGlyAlaGlyGlyProMetGlyGlyProGlnGlnGlnArgP 1327

Qy       6083 CTAAGACACAACACATA------AAGAACATACCTAGTGGAATGTGTCAAAACCAATCGG 6136
                ||            :::         :::|||   |||      |||
Db       1327 roProAsnValGlnValThrProAspGlyMetProMetGlySerGlnGlnGluTrpArgH 1347

Qy       6137 GACTTGCAGTGGCACAAGGGCAGATCCAACTG------------CATGGGCAAGGACATG 6184
                  :::   :::   |||   |||   |||:::                |||   |||
Db       1347 isMetMetMetThrGlnGlnGlnThrGlnMetGlyPheGlyGlyProGlyProGlyGlyP 1367

Qy       6185 CGCAGGGTCAGTCTTTAATTGGACCTACTAATAATAATTTAATG---------------- 6228
```

```
                    |||         |||         |||   |||
Db       1367 roMetArgGlnGlyProGlyGlyPheAsnGlyGlyAsnPheMetProAsnGlyAlaProA 1387

Qy       6229 --TCAACTGCCGGAAGTGTCAGTGCTACTAACGGT---GTCTCTGGCATCAATTTCGTAG 6283
                  |||||||||               |||   :::||||||   |||
Db       1387 snGlyAlaAlaGlySerGlyProAsnAlaGlyGlyMetMetSerGlyProAsnValProG 1407

Qy       6284 GTCCCTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCAGCAGCAGTTAT 6343
                    |||      :::      ||||||            |||||||||
Db       1407 lnMetGlnLeuThrProAlaGlnMetGlnGlnGlnLeuMetArgGlnGlnGlnGln---- 1425

Qy       6344 ATGCTACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCACCAGAGCAACA 6403
                        :::   ||||||||||||||||||            :::||||
Db       1426 -----------GlnGlnGlnGlnGlnGlnGlnHisMetGlyProGlyAlaAlaAsnAsnM 1442

Qy       6404 TG 6405
                ||
Db       1442 et 1442
```

Search completed: November 3, 2005, 01:02:05
Job time : 1116.5 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 22:24:08 ; Search time 1054 Seconds
(without alignments)
6713.387 Million cell updates/sec

Title: US-10-664-859-1
Perfect score: 12037
Sequence: 1 acgagtgcttctcttattat.....aaaaaaaaaaaaaaaaaaa 6909

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10664859/runat_02112005_171807_3627/app_query.fasta_1.7047
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10664859_@CGN_1_1_1295_@runat_02112005_171807_3627 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7280	60.5	1469	1	BCL9_DROME	Q961d9 drosophila
2	338.5	2.8	1530	2	Q67FY3	Q67fy3 brachydanio
3	322.5	2.7	1723	2	Q8T248	Q8t248 dictyosteli
4	320	2.7	1426	1	BCL9_HUMAN	O00512 homo sapien
5	313	2.6	1474	2	Q67FY0	Q67fy0 brachydanio
6	301	2.5	2232	2	Q8IFX6	Q8ifx6 caenorhabdi
7	297.5	2.5	1457	2	Q641L9	Q641l9 mus musculu
8	296.5	2.5	1494	2	Q67FY2	Q67fy2 mus musculu
9	296.5	2.5	1494	2	Q6I7B5	Q6i7b5 mus musculu
10	289.5	2.4	1425	2	Q67FX9	Q67fx9 mus musculu
11	288	2.4	4662	2	Q8ILR9	Q8ilr9 plasmodium
12	281	2.3	1494	2	Q67FY1	Q67fy1 homo sapien
13	280	2.3	1499	2	Q86UU0	Q86uu0 homo sapien
14	279	2.3	1774	2	Q8IAU8	Q8iau8 plasmodium
15	277.5	2.3	1902	2	Q640Q1	Q640q1 mus musculu
16	277.5	2.3	3006	2	Q7YY56	Q7yy56 cryptospori
17	274.5	2.3	1918	2	Q86AF5	Q86af5 dictyosteli
18	270.5	2.2	1596	1	MAM_DROME	P21519 drosophila
19	269.5	2.2	1594	2	Q9V6W7	Q9v6w7 drosophila
20	267.5	2.2	1902	2	Q925Q1	Q925q1 mus musculu
21	261.5	2.2	2206	2	Q7QAV8	Q7qav8 anopheles g
22	261.5	2.2	2235	2	Q8IB09	Q8ib09 plasmodium
23	258.5	2.1	1655	2	Q24754	Q24754 drosophila
24	258.5	2.1	2151	2	Q9VPL6	Q9vpl6 drosophila
25	256.5	2.1	5322	2	Q9VPL9	Q9vpl9 drosophila
26	255.5	2.1	2151	2	Q9Y1L3	Q9y113 drosophila
27	254.5	2.1	1348	2	Q86KE2	Q86ke2 dictyosteli
28	254.5	2.1	2038	2	Q9W3L3	Q9w313 drosophila
29	251.5	2.1	1502	2	Q8IS10	Q8is10 dictyosteli
30	249.5	2.1	2038	1	FSH_DROME	P13709 drosophila
31	249	2.1	1685	2	Q7REY7	Q7rey7 plasmodium
32	248	2.1	3443	2	Q8JZM8	Q8jzm8 mus musculu
33	247.5	2.1	1902	1	SMF1_HUMAN	O14497 homo sapien
34	247.5	2.1	1999	2	Q8NFD6	Q8nfd6 homo sapien
35	247	2.1	1255	2	Q9W3Q5	Q9w3q5 drosophila
36	246.5	2.0	2634	2	Q63KC6	Q63kc6 burkholderi
37	244.5	2.0	2072	2	Q7RSW8	Q7rsw8 plasmodium
38	243.5	2.0	1208	2	Q96TA9	Q96ta9 homo sapien
39	243.5	2.0	1323	2	Q7KSZ0	Q7ksz0 drosophila
40	243.5	2.0	1374	2	Q7YU77	Q7yu77 drosophila
41	243.5	2.0	1376	2	Q9VI63	Q9vi63 drosophila
42	242	2.0	1353	2	Q7RDJ8	Q7rdj8 plasmodium
43	241	2.0	796	2	Q6NRE2	Q6nre2 xenopus lae
44	241	2.0	1606	2	Q63UB0	Q63ub0 burkholderi
45	241	2.0	2280	2	Q9V8E6	Q9v8e6 drosophila

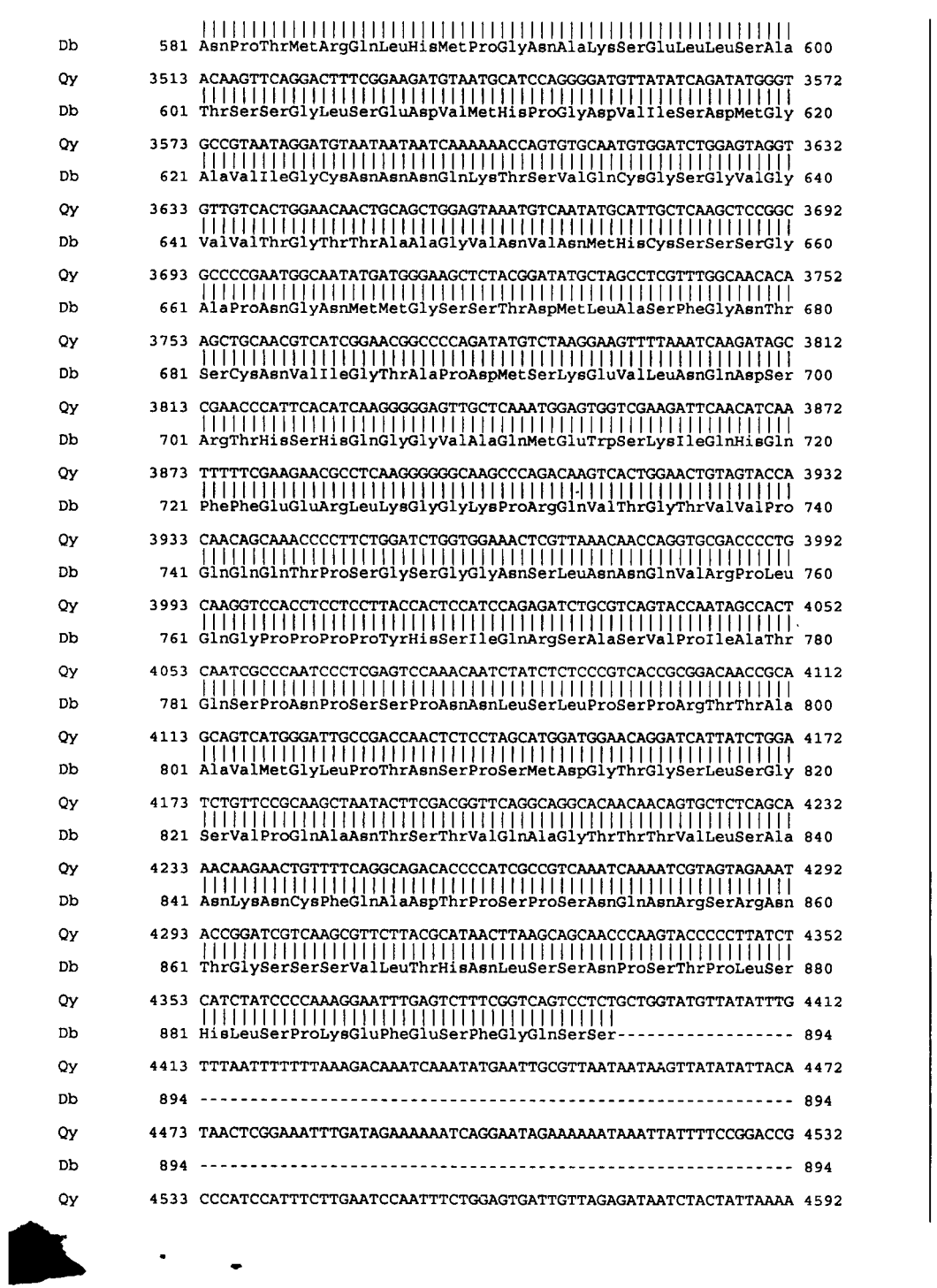
ALIGNMENTS

RESULT 1
BCL9_DROME
ID BCL9_DROME STANDARD; PRT; 1469 AA.
AC Q961D9; Q9V4D2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Bcl-9 homolog (Legless protein).
GN Name=lgs; Synonyms=BCL9; ORFNames=CG2041;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

```
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   GENOME REANNOTATION.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN   [4]
RP   SEQUENCE OF 6-1469 FROM N.A., AND MUTAGENESIS OF GLY-514; LEU-534 AND
RP   ILE-537.
RX   MEDLINE=21952490; PubMed=11955446; DOI=10.1016/S0092-8674(02)00679-7;
RA   Kramps T., Peter O., Brunner E., Nellen D., Froesch B., Chatterjee S.,
RA   Murone M., Zuellig S., Basler K.;
RT   "Wnt/wingless signaling requires BCL9/legless-mediated recruitment of
RT   pygopus to the nuclear beta-catenin-TCF complex.";
RL   Cell 109:47-60(2002).
CC   -!- FUNCTION: Involved in signal transduction through the Wnt pathway.
CC   -!- SUBUNIT: Binds to ARM and PYGO.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically
CC       throughout development.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003844; AAF59345.2; -.
DR   EMBL; AY051651; AAK93075.1; -.
DR   EMBL; AF457205; AAL91368.1; -.
DR   FlyBase; FBgn0039907; lgs.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0030528; F:transcription regulator activity; IPI.
DR   GO; GO:0030177; P:positive regulation of Wnt receptor signali. . .; IPI.
DR   GO; GO:0007367; P:segment polarity determination; IMP.
KW   Developmental protein; Nuclear protein; Segmentation polarity protein;
KW   Wnt signaling pathway.
FT   DOMAIN      511    555       ARM-binding.
FT   DOMAIN     1134   1173       Asn-rich.
FT   DOMAIN     1340   1449       Gln-rich.
FT   DOMAIN     1162   1169       Poly-Asn.
FT   MUTAGEN     514    514       G->E: In allele lgs-21L.
FT   MUTAGEN     534    534       L->F: In allele lgs-17E; segment polarity
FT                                phenotype.
FT   MUTAGEN     537    537       I->K: In allele lgs-17P.
SQ   SEQUENCE   1469 AA;  153759 MW;  5672E01B7200ED08 CRC64;

Alignment Scores:
Pred. No.:              0           Length:         1469
Score:                  7280.00     Matches:        1465
Percent Similarity:     73.08%      Conservative:   1
Best Local Similarity:  73.03%      Mismatches:     3
Query Match:            60.48%      Indels:         537
DB:                     1           Gaps:           5

US-10-664-859-1 (1-6909) x BCL9_DROME (1-1469)

Qy       453 ATGCTCTCGACAACAATGCCCCGCAGTCCAACCCAACAACAGCCGCAACCAAACTCCGAT 512
Db         1 MetLeuSerThrThrMetProArgSerProThrGlnGlnGlnProGlnProAsnSerAsp 20

Qy       513 GCCTCCTCAACAAGTGCATCTGGATCAAATCCTGGAGCAGCGATCGGAAATGGGGACTCG 572
Db        21 AlaSerSerThrSerAlaSerGlySerAsnProGlyAlaAlaIleGlyAsnGlyAspSer 40

Qy       573 GCGGCGAGCAGAAGTTCTCCGAAGACCCTTAATAGCGAACCCTTTTCTACTTTGTCGCCG 632
Db        41 AlaAlaSerArgSerSerProLysThrLeuAsnSerGluProPheSerThrLeu-Ser-- 59

Qy       633 GGTAAGACTTGTATTGATTTCTCTTTGTCCGGAATTATAACAACTTTCTGTGTTTCCAGA 692
Db        60 ----------------------------------------------------------ProAs 61

Qy       693 TCAAATAAAATTGACGCCAGAAGAAGGCACTGAGAAAAGCGGACTATCAACTAGTGATAA 752
Db        61 pGlnIleLysLeuThrProGluGluGlyThrGluLysSerGlyLeuSerThrSerAspLy 81

Qy       753 AGCTGCCACTGGAGGAGCCCCAGGCAGTGGAAATAATCTGCCCGAGGGACAAACTATGCT 812
Db        81 sAlaAlaThrGlyGlyAlaProGlySerGlyAsnAsnLeuProGluGlyGlnThrMetLe 101

Qy       813 AAGGCAGAACTCTACGAGCACAATCAACTCGTGCCTAGTCGCTTCTCCACAAAACTCCAG 872
Db       101 uArgGlnAsnSerThrSerThrIleAsnSerCysLeuValAlaSerProGlnAsnSerSe 121

Qy       873 TGAACACTCGAATAGCAGCAATGTGTCTGCTACAGTGGGCCTTACTCAGATGGTAGATTG 932
Db       121 rGluHisSerAsnSerSerAsnValSerAlaThrValGlyLeuThrGlnMetValAspCy 141

Qy       933 TGACGAGCAATCGAAGAAAAACAAATGTAGTGTGAAGGACGAGGAAGCTGGTAAGACTGC 992
Db       141 sAspGluGlnSerLysLysAsnLysCysSerValLysAspGluGluAla---------- 157

Qy       993 CCTACAAATGGTTTAAAATTTTAAAATGTATTGGCGTTCACCTTTGTTAATCATTTAATT 1052
Db       157 ---------------------------------------------------------- 157

Qy      1053 GTTTTTTTTTTGCTATACTTACAATTTTAGTTTTAAACTTGTAAACTTGACTAAAACTCG 1112
Db       157 ---------------------------------------------------------- 157

Qy      1113 CGAAGCTCGGATCAAAACAGACATTTTCTTGGAACCGTAATTAAGCTCATAAAAATATTA 1172
Db       157 ---------------------------------------------------------- 157

Qy      1173 ATTCATCTTGATGGAATGCATATCATAGATGTACTCAAACATCTCAAGAAAGACCTCAAA 1232
Db       157 ---------------------------------------------------------- 157

Qy      1233 TTGGATCAACTAATTAGTTTGAGAAAAAATTGCTGTACTTTTAAGAATATATTAATTTAA 1292
Db       157 ---------------------------------------------------------- 157
```

```
Qy      1293 AAATTTGCTGAGTGAAATGATATAATAGTCACAATAATTTTTAGTTAAACTGCTAAAGCA 1352
Db       157 ------------------------------------------------------------ 157
Qy      1353 TTTTGAATAGCCGTGCTACGCAGATGCTACTAGACGCGGTGTAAAAGCTAATTTTTATTT 1412
Db       157 ------------------------------------------------------------ 157
Qy      1413 AAAAGCTGTCCTAATATTCCATAACCATTAATGTCCCATTTCAGAAATAAGTTCTAATAA 1472
                                                  ||||||||||||||||
Db       158 ----------------------------------------GluIleSerSerAsnLy 163
Qy      1473 AGCAAAAGGTCAAGCAGCTGGTGGCGGCTGCGAAACAGGTTCTACATCCAGTTTGACTGT 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       163 sAlaLysGlyGlnAlaAlaGlyGlyGlyCysGluThrGlySerThrSerSerLeuThrVa 183
Qy      1533 CAAGGAAGAACCCACCGATGTCTTAGGCAGTTTAGTAAATATGAAAAAAGAAGAAAGAGA 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       183 lLysGluGluProThrAspValLeuGlySerLeuValAsnMetLysLysGluGluArgGl 203
Qy      1593 AAATCATTCGCCAACGATGTCCCCTGTTGGTTTTGGTTCAATTGGTAATGCACAGGACAA 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       203 uAsnHisSerProThrMetSerProValGlyPheGlySerIleGlyAsnAlaGlnAspAs 223
Qy      1653 CTCCGCTACACCGGGTAAGTTTTAAGAGATCCATATAAAGCAAATAACAAGAATTAATGT 1712
              ||||||||||
Db       223 nSerAlaThr---------------------------------------------------- 226
Qy      1713 CAGTTACCAATTTTATTTGATAGTCAAAGAACTACTATAGCGATATCTCCTGCCTTTTAA 1772
Db       226 ------------------------------------------------------------ 226
Qy      1773 TTTTATTTTAATTAGGAAATACGAATATTTCTAATTTGTAAAATAAAATTGATTAATTAA 1832
Db       226 ------------------------------------------------------------ 226
Qy      1833 CTAGAATTTAAAAACCTTTTGAATTAGGACATACCCTTCCAAAAATCAGTAATCATTGGG 1892
Db       226 ------------------------------------------------------------ 226
Qy      1893 AACGAGAGTGTGGTCCCGAAGGAGACTACTATAAAACCTTTTGAGCTATCTGATACTGCA 1952
Db       226 ------------------------------------------------------------ 226
Qy      1953 CGCTACTAAAAATGATTAGTTTAGGAAAATGGGTGTAATTTTGTAGGAAGTTTTCATTTT 2012
Db       226 ------------------------------------------------------------ 226
Qy      2013 AGAAGAAATGTGATTATTTTATTAAACCCCTTCAAGCGGAACTACATTTGTTCTACGATA 2072
Db       226 ------------------------------------------------------------ 226
Qy      2073 TTTTGGAAAAACAAATGGTTAAGTTGGAAAGTGCCTATAAAACAGAATTCCACGGTTTCA 2132
Db       226 ------------------------------------------------------------ 226
Qy      2133 AATACTAACCAGGTTTTTGATTTAATTTTGATTAAATGAGAAATTATCACACTTCAGTTA 2192
Db       226 ------------------------------------------------------------ 226
Qy      2193 AAATGTTTAATTCGATTAAGGTCGGACAATCACAGCAGATTTCCATTTTTGCGTGTATAT 2252
Db       226 ------------------------------------------------------------ 226
Qy      2253 ATAGAAGTCGCCTTCACACTCTTCTGGCGCGCTTCACCACTACGTGGAGTTCCGCCCGCA 2312
Db       226 ------------------------------------------------------------ 226
Qy      2313 GTGATTTATATAGATGATTTACGAGTTATTTAATTTTTTATGGTGTATTTTAATAAATAT 2372
Db       226 ------------------------------------------------------------ 226
```

```
Qy      2373 CTTATTTATTCATTTTACATAGTTAAAATTGAAAGAATTTCAAACGACAGTACCACGGAA 2432
                                                  ||||||||||||||||||||||||
Db       227 ----------Pro--ValLysIleGluArgIleSerAsnAspSerThrThrGlu 240
Qy      2433 AAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATGAGCATGGAAGGCTGCAATCAG 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 LysLysGlySerSerLeuThrMetAsnAsnAspGluMetSerMetGluGlyCysAsnGln 260
Qy      2493 TTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTGCAATTTCGAGCATATTAGTA 2552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 LeuAsnProAspPheIleAsnGluSerLeuAsnAsnProAlaIleSerSerIleLeuVal 280
Qy      2553 AGCGGAGTAGGACCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACT 2612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 SerGlyValGlyProIleProGlyIleGlyValGlyAlaGlyThrGlyAsnLeuLeuThr 300
Qy      2613 GCCAACGCCAATGGAATCTCCTCGGGTAGCAGTAATTGTTTGGATTACATGCAACAGCAA 2672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AlaAsnAlaAsnGlyIleSerSerGlySerSerAsnCysLeuAspTyrMetGlnGlnGln 320
Qy      2673 AATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGC 2732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 AsnHisIlePheValPheSerThrGlnLeuAlaAsnLysGlyAlaGluSerValLeuSer 340
Qy      2733 GGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTG 2792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 GlyGlnPheGlnThrIleIleAlaTyrHisCysThrGlnProAlaThrLysSerPheLeu 360
Qy      2793 GAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTC 2852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GluAspPhePheMetLysAsnProLeuLysIleAsnLysLeuGlnArgHisAsnSerVal 380
Qy      2853 GGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAA 2912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 GlyMetProTrpIleGlyMetGlyGlnValGlyLeuThrProProAsnProValAlaLys 400
Qy      2913 ATAACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAA 2972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 IleThrGlnGlnGlnProHisThrLysThrValGlyLeuLeuLysProGlnPheAsnGln 420
Qy      2973 CATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCT 3032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 HisGluAsnSerLysArgSerThrValSerAlaProSerAsnSerPheValAspGlnSer 440
Qy      3033 GATCCTATGGGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT 3092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 AspProMetGlyAsnGluThrGluLeuMetCysTrpGluGlyGlySerSerAsnThrSer 460
Qy      3093 AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACATCCAGCGAGTCACAG 3152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 ArgSerGlyGlnAsnSerArgAsnHisValAspSerIleSerThrSerSerGluSerGln 480
Qy      3153 GCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGCGAT 3212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AlaIleLysIleLeuGluAlaAlaGlyValAspLeuGlyGlnValThrLysGlySerAsp 500
Qy      3213 CCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAA 3272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 ProGlyLeuThrThrGluAsnAsnIleValSerLeuGlnGlyValLysValProAspGlu 520
Qy      3273 AACCTTACACCACAACAGCGGCAACATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATG 3332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       521 AsnLeuThrProGlnGlnArgGlnHisArgGluGluGlnLeuAlaLysIleLysLysMet 540
Qy      3333 AATCAATTTCTTTTTCCTGAAAATGAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATA 3392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AsnGlnPheLeuPheProGluAsnGluAsnSerValGlyAlaAsnValSerSerGlnIle 560
Qy      3393 ACAAAAATTCCAGGAGATTTAATGATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATA 3452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       561 ThrLysIleProGlyAspLeuMetMetGlyMetSerGlyGlyGlyGlyGlySerIleIle 580
Qy      3453 AATCCGACGATGCGACAACTGCATATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCG 3512
```


Db 581 AsnProThrMetArgGlnLeuHisMetProGlyAsnAlaLysSerGluLeuLeuSerAla 600
Qy 3513 ACAAGTTCAGGACTTTCGGAAGATGTAATGCATCCAGGGGATGTTATATCAGATATGGGT 3572
Db 601 ThrSerSerGlyLeuSerGluAspValMetHisProGlyAspValIleSerAspMetGly 620
Qy 3573 GCCGTAATAGGATGTAATAATAATCAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGT 3632
Db 621 AlaValIleGlyCysAsnAsnAsnGlnLysThrSerValGlnCysGlySerGlyValGly 640
Qy 3633 GTTGTCACTGGAACAACTGCAGCTGGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGC 3692
Db 641 ValValThrGlyThrThrAlaAlaGlyValAsnValAsnMetHisCysSerSerSerGly 660
Qy 3693 GCCCCGAATGGCAATATGATGGGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACA 3752
Db 661 AlaProAsnGlyAsnMetMetGlySerSerThrAspMetLeuAlaSerPheGlyAsnThr 680
Qy 3753 AGCTGCAACGTCATCGGAACGGCCCCAGATATGTCTAAGGAAGTTTTAAATCAAGATAGC 3812
Db 681 SerCysAsnValIleGlyThrAlaProAspMetSerLysGluValLeuAsnGlnAspSer 700
Qy 3813 CGAACCCATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAA 3872
Db 701 ArgThrHisSerHisGlnGlyGlyValAlaGlnMetGluTrpSerLysIleGlnHisGln 720
Qy 3873 TTTTTCGAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACTGTAGTACCA 3932
Db 721 PhePheGluGluArgLeuLysGlyGlyLysProArgGlnValThrGlyThrValValPro 740
Qy 3933 CAACAGCAAACCCCTTCTGGATCTGGTGGAAACTCGTTAAACAACCAGGTGCGACCCCTG 3992
Db 741 GlnGlnGlnThrProSerGlySerGlyGlyAsnSerLeuAsnAsnGlnValArgProLeu 760
Qy 3993 CAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGATCTGCGTCAGTACCAATAGCCACT 4052
Db 761 GlnGlyProProProProTyrHisSerIleGlnArgSerAlaSerValProIleAlaThr 780
Qy 4053 CAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTCCCGTCACCGCGGACAACCGCA 4112
Db 781 GlnSerProAsnProSerSerProAsnAsnLeuSerLeuProSerProArgThrThrAla 800
Qy 4113 GCAGTCATGGGATTGCCGACCAACTCTCCTAGCATGGATGGAACAGGATCATTATCTGGA 4172
Db 801 AlaValMetGlyLeuProThrAsnSerProSerMetAspGlyThrGlySerLeuSerGly 820
Qy 4173 TCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGCACAACAACAGTGCTCTCAGCA 4232
Db 821 SerValProGlnAlaAsnThrSerThrValGlnAlaGlyThrThrThrValLeuSerAla 840
Qy 4233 AACAAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCAAATCAAAATCGTAGTAGAAAT 4292
Db 841 AsnLysAsnCysPheGlnAlaAspThrProSerProSerAsnGlnAsnArgSerArgAsn 860
Qy 4293 ACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGCAGCAACCCAAGTACCCCCTTATCT 4352
Db 861 ThrGlySerSerSerValLeuThrHisAsnLeuSerSerAsnProSerThrProLeuSer 880
Qy 4353 CATCTATCCCCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTG 4412
Db 881 HisLeuSerProLysGluPheGluSerPheGlyGlnSerSer---------- 894
Qy 4413 TTTAATTTTTTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACA 4472
Db 894 ---------- 894
Qy 4473 TAACTCGGAAATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCG 4532
Db 894 ---------- 894
Qy 4533 CCCATCCATTTCTTGAATCCAATTTCTGGAGTGATTGTTAGAGATAATCTACTATTAAAA 4592
Db 894 ---------- 894
Qy 4593 TTAAACACGAAAATTCATATCCGTTAATTGAAAATCACTATTGTTTAATAAGAAATTAAA 4652
Db 894 ---------- 894
Qy 4653 AATATGTTTATTATAATATTTCTACAGGTGATAACATGAAAAGTAGGCGACCAAGCCCAC 4712
Db 895 ----------Ala--GlyAspAsnMetLysSerArgArgProSerProG 907
Qy 4713 AGGGTCAGCGGTCACCAGTAAATAGTCTAATAGAGGCAAATAAAGATGTACGATTTGCTG 4772
Db 907 lnGlyGlnArgSerProValAsnSerLeuIleGluAlaAsnLysAspValArgPheAlaA 927
Qy 4773 CATCCAGTCCTGGTTTTAACCCGCATCCACATATGCAAAGCAATTCAAATTCAGCATTAA 4832
Db 927 laSerSerProGlyPheAsnProHisProHisMetGlnSerAsnSerAsnSerAlaLeuA 947
Qy 4833 ACGCCTATAAAATGGGCTCTACCAATATACAGATGGAGGTAAATATTTAAATATTTTATT 4892
Db 947 snAlaTyrLysMetGlySerThrAsnIleGlnMetGlu-Arg---------- 960
Qy 4893 TAACGTTTTTGTGTTAATTTATCTTCTTTTTCAGCGTCAAGCATCAGCGCAAGGTGGATC 4952
Db 961 ----------GlnAlaSerAlaGlnGlyGlySe 968
Qy 4953 CGTACAATTTAGTCGGCGCTCCGATAATATTCCGCTAAATCCCAATAGTGGCAATCGGCC 5012
Db 968 rValGlnPheSerArgArgSerAspAsnIleProLeuAsnProAsnSerGlyAsnArgPr 988
Qy 5013 GCCACCAAACAAGATGACCCAAAACTTCGATCCAATCTCTTCTTTGGCACAAATGTCCCA 5072
Db 988 oProProAsnLysMetThrGlnAsnPheAspProIleSerSerLeuAlaGlnMetSerGl 1008
Qy 5073 ACAACTAACAAGTTGCGTGTCCAGCATGGGTAGTCCAGCCGGAACTGGTGGTATGACGAT 5132
Db 1008 nGlnLeuThrSerCysValSerSerMetGlySerProAlaGlyThrGlyGlyMetThrMe 1028
Qy 5133 GATGGGGGGTCCGGGACCGTCCGACATCAATATTGAGCATGGAATAATTTCGGGACTAGA 5192
Db 1028 tMetGlyGlyProGlyProSerAspIleAsnIleGluHisGlyIleIleSerGlyLeuAs 1048
Qy 5193 TGGATCAGGAATAGATACCATAAATCAAAATAACTGTCATTCAATGAATGTCGTAATGAA 5252
Db 1048 pGlySerGlyIleAspThrIleAsnGlnAsnAsnCysHisSerMetAsnValValMetAs 1068
Qy 5253 CTCAATGGGTCCCCGAATGCTGAATCCTAAAATGTGCGTAGCAGGCGGTCCAAATGGACC 5312
Db 1068 nSerMetGlyProArgMetLeuAsnProLysMetCysValAlaGlyGlyProAsnGlyPr 1088
Qy 5313 GCCTGGCTTTAATCCTAATTCCCCCAATGGTGGATTAAGAGAGAATTCCATAGGGTCTGG 5372
Db 1088 oProGlyPheAsnProAsnSerProAsnGlyGlyLeuArgGluAsnSerIleGlySerGl 1108
Qy 5373 CTGTGGCTCAGCAAACTCTTCAAACTTTCAAGGGGTTGTTCCACCTGGTGCCAGAATGAT 5432
Db 1108 yCysGlySerAlaAsnSerSerAsnPheGlnGlyValValProProGlyAlaArgMetMe 1128
Qy 5433 GGGTCGAATGCCAGTCAATTTTGGTTCGAATTTCAATCCGAATATTCAGGTAAAGGCGAG 5492
Db 1128 tGlyArgMetProValAsnPheGlySerAsnPheAsnProAsnIleGlnValLysAlaSe 1148
Qy 5493 TACCCCAAACACCATACAATACATGCCAGTAAGGGCACAGAACGCCAACAACAATAACAA 5552
Db 1148 rThrProAsnThrIleGlnTyrMetProValArgAlaGlnAsnAlaAsnAsnAsnAsAs 1168
Qy 5553 CAATGGAGCTAATAATGTGCGAATGCCACCTAGTCTGGAATTTTTGCAGAGGTACGCTAA 5612
Db 1168 nAsnGlyAlaAsnAsnValArgMetProProSerLeuGluPheLeuGlnArgTyrAlaAs 1188
Qy 5613 CCCTCAAATGGGTGCTGTAGGCAATGGGTCGCCAATATGCCCACCATCAGCCAGCGACGG 5672
Db 1188 nProGlnMetGlyAlaValGlyAsnGlySerProIleCysProProSerAlaSerAspGl 1208

```
Qy       5673 TACTCCTGGAATGCCAGGATTGATGGCGGGACCAGGAGCCGGAGGTATGCTAATGAATTC 5732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 yThrProGlyMetProGlyLeuMetAlaGlyProGlyAlaGlyGlyMetLeuMetAsnSe 1228

Qy       5733 TTCCGGAGAGCAACACCAGAACAAGATCACAAACAATCCTGGGGCAAGCAATGGTATTAA 5792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 rSerGlyGluGlnHisGlnAsnLysIleThrAsnAsnProGlyAlaSerAsnGlyIleAs 1248

Qy       5793 CTTCTTTCAGAATTGCAATCAAATGTCTATTGTTGACGAAGAGGGTGGATTACCCGGCCA 5852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 nPhePheGlnAsnCysAsnGlnMetSerIleValAspGluGluGlyGlyLeuProGlyHi 1268

Qy       5853 TGACGGATCAATGAATATTGGTCAACCATCTATGATAAGGGGCATGCGTCCACATGCCAT 5912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 sAspGlySerMetAsnIleGlyGlnProSerMetIleArgGlyMetArgProHisAlaMe 1288

Qy       5913 GCGGCCAAATGTAATGGGTGCGCGGATGCCACCCGTTAACAGGCAAATTCAGTTTGCACA 5972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 tArgProAsnValMetGlyAlaArgMetProProValAsnArgGlnIleGlnPheAlaGl 1308

Qy       5973 GTCATCGGATGGTATTGACTGTGTCGGGGATCCGTCATCATTTTTCACTAACGCTTCCTG 6032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 nSerSerAspGlyIleAspCysValGlyAspProSerSerPhePheThrAsnAlaSerCy 1328

Qy       6033 CAACAGCGCTGGACCACACATGTTTGGATCAGCACAACAGGCCAATCAGCCTAAGACACA 6092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 sAsnSerAlaGlyProHisMetPheGlySerAlaGlnGlnAlaAsnGlnProLysThrGl 1348

Qy       6093 ACACATAAAGAACATACCTAGTGGAATGTGTCAAAACCAATCGGGACTTGCAGTGGCACA 6152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 nHisIleLysAsnIleProSerGlyMetCysGlnAsnGlnSerGlyLeuAlaValAlaGl 1368

Qy       6153 AGGGCAGATCCAACTGCATGGGCAAGGACATGCGCAGGGTCAGTCTTTAATTGGACCTAC 6212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 nGlyGlnIleGlnLeuHisGlyGlnGlyHisAlaGlnGlyGlnSerLeuIleGlyProTh 1388

Qy       6213 TAATAATAATTTAATGTCAACTGCCGGAAGTGTCAGTGCTACTAACGGTGTCTCTGGCAT 6272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 rAsnAsnAsnLeuMetSerThrAlaGlySerValSerAlaThrAsnGlyValSerGlyIl 1408

Qy       6273 CAATTTCGTAGGTCCCTCTTCTACGGACCTGAAGTATGCCCAGCAATATCATAGTTTTCA 6332
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 eAsnPheValGlyProSerSerThrAspLeuLysTyrAlaGlnGlnTyrHisSerPheGl 1428

Qy       6333 GCAGCAGTTATATGCTACCAACACCAGAAGTCAACAACAACAGCATATGCACCAGCAGCA 6392
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1428 nGlnGlnLeuTyrAlaThrAsnThrArgSerGlnGlnGlnGlnHisMetHisGlnGlnHi 1448

Qy       6393 CCAGAGCAACATGATAACAATGCCGCCGAATTTATCACCAAATCCAACGTTCTTTGTCAA 6452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 sGlnSerAsnMetIleThrMetProProAsnLeuSerProAsnProThrPhePheValAs 1468

Qy       6453 CAAA 6456
             ||||
Db       1468 nLys 1469
```

RESULT 2
Q67FY3
ID Q67FY3 PRELIMINARY; PRT; 1530 AA.
AC Q67FY3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Bcl9-2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Brembeck F.H., Schwarz-Romond T., Bakkers J., Wilhelm S.,
RA Hammerschmidt M., Birchmeier W.;
RT "Essential role of BCL9-2 in the switch between [beta]-catenin's
RT adhesive and transcriptional functions.";
RL Genes Dev. 18:0-0(2004).
DR EMBL; AY296057; AAQ62695.1; -.
SQ SEQUENCE 1530 AA; 159872 MW; C29FEC9433ED28C0 CRC64;

Alignment Scores:

Pred. No.:	9.93e-14	Length:	1530
Score:	338.50	Matches:	308
Percent Similarity:	30.82%	Conservative:	171
Best Local Similarity:	19.82%	Mismatches:	554
Query Match:	2.81%	Indels:	522
DB:	2	Gaps:	71

US-10-664-859-1 (1-6909) x Q67FY3 (1-1530)

```
Qy       2565 CCAATACCCGGAATCGGAGTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAAT 2624
             |||      |||          :::   |||||::::::
Db        153 ProSerGlnGlyHisSerGlySerSerThrThrGlyHisValSerAspProGlyGlyPro 172

Qy       2625 GGAATCTCCTCGGGTAGCAGT-----------AATTGTTTGGATTACATGCAACAGCAA 2672
             |||:::   ||||||                 |||                  |||
Db        173 GlyLeuGlySerGlyHisGlyProGlyIleArgThrAspLeuHisSerArgProProGln 192

Qy       2673 AATCACATATTCGTGTTTTCAACTCAGCTGGCCAACAAAGGGGCCGAATCAGTTTTAAGC 2732
                 ::::::||||||:::|||   |||||||||      ||||||:::|||:::
Db        193 GlnValValTyrValPheThrThrSerLeuAlaAsnSerAlaAlaGluAlaValMetHis 212

Qy       2733 GGTCAATTTCAAACTATTATTGCGTATCACTGCACTCAGCCTGCTACAAAAAGCTTCCTG 2792
             |||          :::|||:::   ||||||      |||         ::::::   |||
Db        213 GlyHisThrAspSerIleLeuLeuTyrHis---GlnGlnAsnValProArgThrLysLeu 231

Qy       2793 GAAGACTTTTTTATGAAAAACCCTTTAAAGATTAACAAGTTACAGCGGCACAATTCCGTC 2852
             :::                  |||::::::   |||       |||   |||
Db        232 AspGln------SerThrGlyValGlyLysValSerAsnLeuAlaGluHisIleSerSer 249

Qy       2853 GGTATGCCATGGATAGGCATGGGGCAGGTTGGACTAACTCCTCCTAATCCTGTAGCCAAA 2912
                                                 :::||||||            |||
Db        250 SerHis-----------------------------SerProProIleGlyThrProLys 259

Qy       2913 ATAACACAACAGCAGCCACATACAAAGACCGTAGGCCTATTGAAACCCCAATTCAATCAA 2972
                           |||         :::||||||   :::
Db        260 SerGlnSerGlyThrProArgProAlaSerValGlyGlyVal------------------ 273

Qy       2973 CATGAAAACAGCAAACGTAGTACTGTAAGCGCGCCTAGCAACTCTTTTGTCGACCAGTCT 3032
                                     |||   ||||||:::::              |||
Db        274 ---------GlyHisLeuProGlyThrSerThrProSerSerThrGlyHisProAspSer 290

Qy       3033 GATCCTATGGGCAACGAAACTGAATTGATGTGCTGGGAAGGCGGATCCTCAAACACCAGT 3092
             :::|||                                |||||   :::
Db        291 GluProAlaGlnThrHis-----------------ArgGlyGlyGlyThr--------- 301

Qy       3093 AGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATCAGTACA---TCCAGCGAGTCA 3149
                |||      |||   |||:::   |||   ::::::          :::|||
Db        302 ---SerSerAsnAsnGlyArgSerAlaValHisThrLeuGlyLeuGlyAsnSerGlyPro 320

Qy       3150 CAGGCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGACAGGTCACAAAAGGAAGC 3209
             |||::::::   :::            |||||||||
Db        321 GlnSerValGlyValSerGlyThrGluGlyValAspArgProGlyAlaIleProHisHis 340

Qy       3210 GATCCTGGCCTGACAACTGAAAACAACATTGTATCACTGCAAGGAGTTAAGGTTCCAGAC 3269
                ||::::::      :::          |||      ::::::      :::
Db        341 GlyAlaGlyValSerProSerThrSerProSerValLeuSerAlaLeuArgGlnSerGlu 360

Qy       3270 -----------------------------GAAAACCTTACACCACAACAGCGGCAACAT 3299
                                           :::   |||:::   :::|||   :::|||
Db        361 LeuGlyGlnArgValGlyProGlyAsnThrAspGlyLeuSerLysGluGlnLeuGluHis 380

Qy       3300 CGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTTCCTGAAAATGAG 3359
             ||||||      |||      ::::::   :::   :::   |||
Db        381 ArgGluArgSerLeuGlnThrLeuArgAspIleGluArgLeuLeuLeuArgSer------ 398
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 2, 2005, 23:46:44 ; Search time 206 Seconds
(without alignments)
6453.996 Million cell updates/sec

Title: US-10-664-859-1
Perfect score: 12037
Sequence: 1 acgagtgcttctcttattat.....aaaaaaaaaaaaaaaaaaa 6909

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10664859/runat_02112005_171808_3665/app_query.fasta_1.7047
-DB=PIR -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10664859_@CGN_1_1_338_@runat_02112005_171808_3665 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	301	2.5	2232	2	T34434	hypothetical prote
2	270.5	2.2	1596	2	A33106	neurogenic locus m
3	258.5	2.1	1655	2	T13998	gene mastermind pr
4	249.5	2.1	2038	2	A43742	female sterile hom
5	238.5	2.0	2715	2	T13049	eyelid - fruit fly
6	233	1.9	3190	2	T13828	CREB-binding prote
7	231.5	1.9	1844	2	D71612	hypothetical prote
8	228	1.9	1424	2	T03851	thyroid hormone re
9	222.5	1.8	1063	2	D86731	hypothetical prote
10	222.5	1.8	1459	2	T32271	hypothetical prote
11	219	1.8	1457	2	T14577	protein kinase Yak
12	216.5	1.8	1061	2	A57620	steroid receptor c
13	215	1.8	786	2	T16509	hypothetical prote
14	214.5	1.8	1142	2	T00022	B120 protein - hum
15	213	1.8	925	2	T19361	hypothetical prote
16	207	1.7	1441	2	B86807	hypothetical prote
17	206	1.7	1088	2	H96747	unknown protein T1
18	205.5	1.7	959	2	B44402	nuclear pore compl
19	205	1.7	1237	2	D71850	probable outer mem
20	204	1.7	1584	2	T18276	protein-tyrosine k
21	203.5	1.7	13288	2	T03099	mucin, submaxillar
22	201.5	1.7	1032	2	T34433	hypothetical prote
23	201.5	1.7	1436	2	D71618	hypothetical prote
24	201	1.7	1778	2	T50074	probable nucleopor
25	200.5	1.7	907	2	A45560	sporozoite surface
26	198	1.6	1230	2	E64664	outer membrane pro
27	197	1.6	1532	2	A61262	collagen alpha 1(X
28	195.5	1.6	3968	2	A44265	trithorax homolog
29	194.5	1.6	1436	2	S57238	forked protein 5.4
30	194.5	1.6	1449	2	S57237	forked protein 5.6
31	193	1.6	1113	2	S28925	nuclear pore compl
32	193	1.6	2468	2	A83412	hypothetical prote
33	193	1.6	2893	2	A64556	toxin-like outer m
34	192	1.6	2703	1	A24420	notch protein - fr
35	191.5	1.6	2902	2	C71953	toxin-like outer m
36	190	1.6	1307	2	T25563	hypothetical prote
37	189	1.6	1217	2	S52714	sericin1B - silkwo
38	188.5	1.6	2150	2	S71629	sensory transducti
39	188	1.6	1072	2	A86827	hypothetical prote
40	188	1.6	1296	2	T13936	collar protein iso
41	188	1.6	2027	2	S60123	hypothetical prote
42	188	1.6	2056	2	G88564	protein R10E11.1 [
43	186	1.5	1910	2	AF0394	probable adhesin h
44	185.5	1.5	1858	2	T18273	1-phosphatidylinos
45	185.5	1.5	3507	2	T34513	hypothetical prote

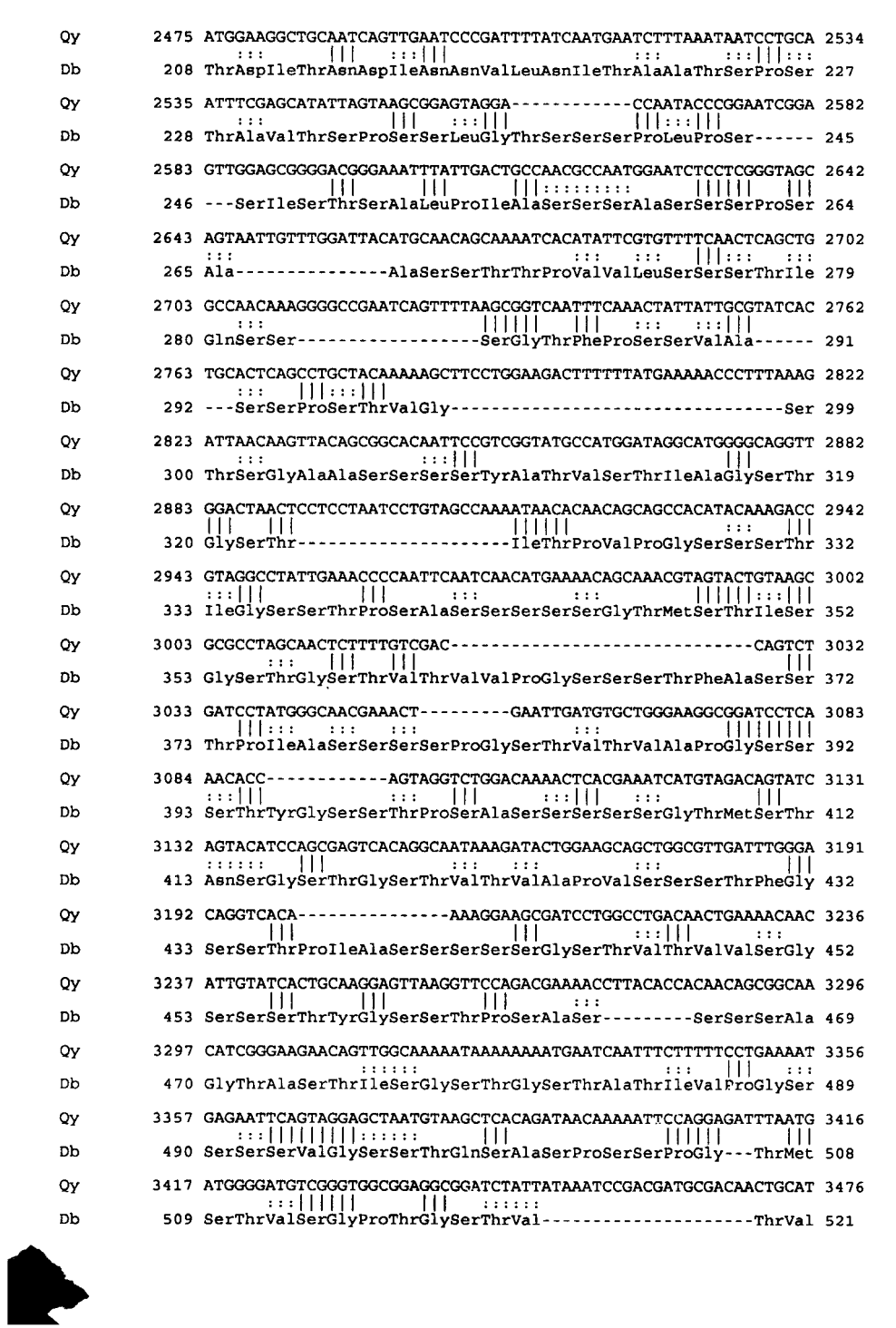
ALIGNMENTS

RESULT 1
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34434
R;Geisel, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid K06A9.
A;Reference number: Z21525
A;Accession: T34434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2232 <GEI>
A;Cross-references: UNIPROT:Q8IFX6; EMBL:U80846; PIDN:AAC70890.1; GSPDB:GN00028; CESP:K
A;Experimental source: strain Bristol N2; clone K06A9
C;Genetics:
A;Gene: CESP:K06A9.1a
A;Map position: X
A;Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/

Alignment Scores:
Pred. No.: 6.48e-12 Length: 2232
Score: 301.00 Matches: 297
Percent Similarity: 33.17% Conservative: 187
Best Local Similarity: 20.36% Mismatches: 599
Query Match: 2.50% Indels: 378
DB: 2 Gaps: 62

US-10-664-859-1 (1-6909) x T34434 (1-2232)

```
Qy       2355 GTGTATTTTAATAAATATCTTATTTATTCATTTTACATAGTTAAAATTGAAAGAATTTCA 2414
                 |||         ||||||:::|||   |||                      :::
Db        174 ValGlnThrLysLysTyrValIleValSerLeu-----------------AlaMetAsp 187

Qy       2415 AACGACAGTACCACGGAAAAAAAAGGATCGTCCTTGACAATGAATAATGACGAAATGAGC 2474
                  :::   |||:::|||   |||         |||   |||            :::
Db        188 GlnGlyAsnMetThrGlnLysTyrGlyAspLeuAsnThrHisAsnIleGlyAsnAsnAla 207
```


Qy 2475 ATGGAAGGCTGCAATCAGTTGAATCCCGATTTTATCAATGAATCTTTAAATAATCCTGCA 2534
Db 208 ThrAspIleThrAsnAspIleAsnAsnValLeuAsnIleThrAlaAlaThrSerProSer 227
Qy 2535 ATTTCGAGCATATTAGTAAGCGGAGTAGGA----------CCAATACCCGGAATCGGA 2582
Db 228 ThrAlaValThrSerProSerSerLeuGlyThrSerSerSerProLeuProSer------ 245
Qy 2583 GTTGGAGCGGGGACGGGAAATTTATTGACTGCCAACGCCAATGGAATCTCCTCGGGTAGC 2642
Db 246 ---SerIleSerThrSerAlaLeuProIleAlaSerSerSerAlaSerSerSerProSer 264
Qy 2643 AGTAATTGTTTGGATTACATGCAACAGCAAAATCACATATTCGTGTTTTCAACTCAGCTG 2702
Db 265 Ala--------------AlaSerSerThrThrProValValLeuSerSerSerThrIle 279
Qy 2703 GCCAACAAAGGGGCCGAATCAGTTTTAAGCGGTCAATTTCAAACTATTATTGCGTATCAC 2762
Db 280 GlnSerSer----------------SerGlyThrPheProSerSerValAla------ 291
Qy 2763 TGCACTCAGCCTGCTACAAAAAGCTTCCTGGAAGACTTTTTTATGAAAAACCCTTTAAAG 2822
Db 292 ---SerSerProSerThrValGly--------------------------------Ser 299
Qy 2823 ATTAACAAGTTACAGCGGCACAATTCCGTCGGTATGCCATGGATAGGCATGGGGCAGGTT 2882
Db 300 ThrSerGlyAlaAlaSerSerSerSerTyrAlaThrValSerThrIleAlaGlySerThr 319
Qy 2883 GGACTAACTCCTCCTAATCCTGTAGCCAAAATAACACAACAGCAGCCACATACAAAGACC 2942
Db 320 GlySerThr--------------------IleThrProValProGlySerSerSerThr 332
Qy 2943 GTAGGCCTATTGAAACCCCAATTCAATCAACATGAAAACAGCAAACGTAGTACTGTAAGC 3002
Db 333 IleGlySerSerThrProSerAlaSerSerSerSerSerGlyThrMetSerThrIleSer 352
Qy 3003 GCGCCTAGCAACTCTTTTGTCGAC----------CAGTCT 3032
Db 353 GlySerThrGlySerThrValThrValValProGlySerSerSerThrPheAlaSerSer 372
Qy 3033 GATCCTATGGGCAACGAAACT---------GAATTGATGTGCTGGGAAGGCGGATCCTCA 3083
Db 373 ThrProIleAlaSerSerSerSerProGlySerThrValThrValAlaProGlySerSer 392
Qy 3084 AACACC------------AGTAGGTCTGGACAAAACTCACGAAATCATGTAGACAGTATC 3131
Db 393 SerThrTyrGlySerSerThrProSerAlaSerSerSerSerSerGlyThrMetSerThr 412
Qy 3132 AGTACATCCAGCGAGTCACAGGCAATAAAGATACTGGAAGCAGCTGGCGTTGATTTGGGA 3191
Db 413 AsnSerGlySerThrGlySerThrValThrValAlaProValSerSerSerThrPheGly 432
Qy 3192 CAGGTCACA--------------AAAGGAAGCGATCCTGGCCTGACAACTGAAAACAAC 3236
Db 433 SerSerThrProIleAlaSerSerSerSerSerGlySerThrValThrValSerGly 452
Qy 3237 ATTGTATCACTGCAAGGAGTTAAGGTTCCAGACGAAAACCTTACACCACAACAGCGGCAA 3296
Db 453 SerSerSerThrTyrGlySerSerThrProSerAlaSer---------SerSerSerAla 469
Qy 3297 CATCGGGAAGAACAGTTGGCAAAAATAAAAAAAATGAATCAATTTCTTTTTCCTGAAAAT 3356
Db 470 GlyThrAlaSerThrIleSerGlySerThrGlySerThrAlaThrIleValProGlySer 489
Qy 3357 GAGAATTCAGTAGGAGCTAATGTAAGCTCACAGATAACAAAAATTCCAGGAGATTTAATG 3416
Db 490 SerSerSerValGlySerSerThrGlnSerAlaSerProSerSerProGly---ThrMet 508
Qy 3417 ATGGGGATGTCGGGTGGCGGAGGCGGATCTATTATAAATCCGACGATGCGACAACTGCAT 3476
Db 509 SerThrValSerGlyProThrGlySerThrVal--------------------ThrVal 521
Qy 3477 ATGCCAGGTAACGCCAAATCGGAGCTCTTATCGGCGACAAGTTCAGGACTTTCGGAAGAT 3536
Db 522 ValProGlySerSerThrSerProAlaProSerSerSerPro---------- 535
Qy 3537 GTAATGCATCCAGGGGATGTTATATCAGATATGGGTGCCGTAATAGGATGTAATAATAAT 3596
Db 536 ------AsnProSerSerSerProAlaSerThrGlySerThrIleThrIleSerGlySer 553
Qy 3597 CAAAAAACCAGTGTGCAATGTGGATCTGGAGTAGGTGTTGTCACTGGAACAACTGCAGCT 3656
Db 554 SerSerIleIleValSerThrValSerGly---SerThrValSerGlySerThr------ 570
Qy 3657 GGAGTAAATGTCAATATGCATTGCTCAAGCTCCGGCGCCCCG---AATGGCAATATGATG 3713
Db 571 GlyThrSerGlnSerThrLeuAlaSerSerThrAlaThrProGlySerSerSerThrVal 590
Qy 3714 GGAAGCTCTACGGATATGCTAGCCTCGTTTGGCAACACAAGCTGCAACGTCATCGGAACG 3773
Db 591 ProSerSerSerSerProGlnProSerSerGlnSerProAlaProAsnThrGlySerThr 610
Qy 3774 GCCCCAGATATGTCTAAGGAA--------------GTTTTAAATCAAGATAGCCGAACC 3818
Db 611 ThrProSerGlnThrSerSerGlnSerProSerProSerMetAsnProSerSerSerThr 630
Qy 3819 CATTCACATCAAGGGGGAGTTGCTCAAATGGAGTGGTCGAAGATTCAACATCAATTTTTC 3878
Db 631 ProThr-----------------------------GlySerSerGlnSerThrIleThr 640
Qy 3879 GAAGAACGCCTCAAGGGGGGCAAGCCCAGACAAGTCACTGGAACT-------------- 3923
Db 641 ProGluGlySerThrAlaSerSerProThrGlySerThrGlySerThrPheSerValAla 660
Qy 3924 ------GTAGTACCACAACAGCAAACCCCTTCTGGATCT------GGTGGAAACTCGTTA 3971
Db 661 ThrGluValThrSerGlnSerThrValProSerGlySerSerLeuGlyThrGlnSerThr 680
Qy 3972 AACAACCAGGTGCGACCCCTGCAAGGTCCACCTCCTCCTTACCACTCCATCCAGAGATCT 4031
Db 681 AsnSerSer---------------------ProSerProSerSerLeuSerProSerThr 693
Qy 4032 GCGTCAGTACCAATAGCCACTCAATCGCCCAATCCCTCGAGTCCAAACAATCTATCTCTC 4091
Db 694 SerGlyMetSerThrLeuThrSerGluProSerProSerSerThrGlnSerSerGlyAla 713
Qy 4092 CCGTCACCGCGGACAACCGCAGCAGTCATGGGATTGCCGACCAACTCTCCTAGCATGGAT 4151
Db 714 GlnSerThrLeuThrThr-----------------ProSerProAsnProSer---Gln 726
Qy 4152 GGAACAGGATCATTATCTGGATCTGTTCCGCAAGCTAATACTTCGACGGTTCAGGCAGGC 4211
Db 727 SerThrSerSerLeuGluSerSerThrSerGlyAlaThrThrSerSerGlySerAlaGly 746
Qy 4212 ACAACAACAGTGCTCTCAGCAAACAAGAACTGTTTTCAGGCAGACACCCCATCGCCGTCA 4271
Db 747 ThrThrMetThr--SerProSer 753
Qy 4272 AATCAAAATCGTAGTAGAAATACCGGATCGTCAAGCGTTCTTACGCATAACTTAAGCAGC 4331
Db 754 ---------GlnSerSerSerValGlySerSerGlnGlySerThrSerProAlaAlaSer 770
Qy 4332 AACCCAAGT----------ACCCCCTTATCTCATCTATCC 4361
Db 771 ThrThrSerGlyGluMetThrSerGlnGlySerThrGlnThrProGlySerSerValSer 790
Qy 4362 CCAAAGGAATTTGAGTCTTTCGGTCAGTCCTCTGCTGGTATGTTATATTTGTTTAATTTT 4421
Db 791 -------------------------ThrSerAlaAlaIleLeu--------------- 796
Qy 4422 TTTAAAGACAAATCAAATATGAATTGCGTTAATAATAAGTTATATATTACATAACTCGGA 4481
Db 797 ---------ThrSerThrGlnGlnSerValSerThrAsnSer-----------ProGly 809
Qy 4482 AATTTGATAGAAAAAATCAGGAATAGAAAAAATAAATTATTTTCCGGACCGCCCATCCAT 4541